(Founded 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

EDITED BY

CARL MURCHISON

Gordon Allport
Harvard University

Arthur G. Bills
University of Cincinnati

E. A. Bott
University of Toronto

Leonard Carmichael
Smithsonian Institution

Herbert S. Conrad
U. S. Office of Education

Elmer K. Culler
University of Rochester

John F. Dashiell
University of North Carolina

Roy M. Dorcus
University of California at Los Angeles

Horace B. English
Ohio State University

Franklin Fearing
University of California at Los Angeles

Frank A. Geldard
University of Virginia

Clarence H. Graham
Columbia University

Ernest R. Hilgard
Stanford University

Heinrich Klüver
University of Chicago

Carney Landis
Columbia University

Norman L. Munn
Bowdoin University

John Paul Nafe
Washington University

Donald G. Paterson
University of Minnesota

Lee Edward Travis
University of Southern California

Robert C. Tryon
University of California

Michael J. Zigler
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 50

1960

Copyright, 1960, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Barrrows, J. 205
Bartley, S. Howard 53, 315, 323
Bell, D. M. 145
Berringer, Richard D. 75
Bronzaft, Arline 193, 279
Bruce, Roselyn Sherman 15
Crowder, William F. 287
DeHardt, Doris C. 53, 315
Dykman, Roscoe A. 105, 227
Edgington, Eugene S. 399
Edwards, Allan E. 139
Elkind, David 111
Flores, Ivan 3
Galbrecht, Charles R. 227
Gantt, W. Horsley 105
Gladin, Leo L. 59
Gogel, Walter C. 119, 179, 257
Goldfard, L. 85
Goodchilds, Jacqueline D. 75
Gorfein, D. 205
Gray, Horace 383
Hawkes, Glenn R. 303
Hayes, Roslyn 193, 279
Israel, M. L. 43
Jensen, Milton B. 211
Kindrick, T. 205
Koltuv, Myron 193, 279
Krampen, Martin 271
Krinsky, L. W. 85
Lehner, George F. J. 333
Lebo, Dell 15, 197
Leland, Q. 205
Levy, Leo 39
Marquart, Dorothy Irene 291
Mayzner, M. S. 339, 349
McAvoy, M. E. 205
Meade, Robert D. 173
Minas, J. Sayer 133
Nash, Franklin A., Jr. 287
Nelson, Thomas M. 323
O'Brien, Cyril C. 377
Peters, John E. 227
Plutchik, Robert 153
Reitan, Ralph M. 355
Rolo, A. 85
Schatz, Joel 253
Scodel, Alvin 133
Shor, Donald E. 253
Spigel, Irwin M. 209
Tresselt, M. E. 339, 349
Welch, Livingston 193, 279
Wells, H. M. 145
Ziller, Robert C. 75

Edn Research
3.9.70
J- 329

TABLE OF CONTENTS

Methods for comparing the legibility of printed numerals 3
By Ivan Flores

Projective methods recommended for use with the blind 15
By Dell Lebo and Roselyn Sherman Bruce

Studies in conformity behavior: A methodological note 39
By Leo Levy

Variably blurred prompting: I. Methodology and application to the analysis of paired-associate learning 43
By M. L. Israel

A further factor in determining nearness as a function of lateral orientation in pictures 53
By S. Howard Bartley and Doris C. DeHardt

The return of the homunculus 59
By Leo L. Gladin

The minority newcomer in open and closed groups 75
By Robert C. Ziller, Richard D. Behringer, and Jacqueline D. Goodchilds

LSD as an adjunct to psychotherapy with alcoholics 85
By A. Rolo, L. W. Krinsky, and L. Goldfarb

A case of experimental neurosis and recovery in relation to the orienting response 105
By Roscoe A. Dykman and W. Horsley Gantt

Interviewing children in a school setting 111
By David Elkind

Perceived frontal size as a determiner of perceived binocular depth . . . 119
By Walter C. Gogel

The behavior of prisoners in a "prisoner's dilemma" game 133
By Alvin Scodel and J. Sayer Minas

Subliminal tachistoscopic perception as a function of threshold method . . 139
By Allan E. Edwards

Binocular rivalry of parental figures in the humascope 145
By H. M. Wells and D. M. Bell

The multifactor-analytic theory of emotion 153
By Robert Plutchik

Time estimates as affected by need tension and rate of progress . . . 173
By Robert D. Meade

The perception of shape from binocular disparity cues 179
By Walter C. Gogel

Relationship between *PGR* and measures of extraversion, ascendance, and neuroticism 193
By Arline Bronzaft, Roslyn Hayes, Livingston Welch, and Myron Koltuv

The development and employment of VTAT's or pictureless TAT's . . . 197
By Dell Lebo

Cognitive dissinance and yielding behavior 205
By D. Gorfein, T. Kindrick, Q. Leland, M. E. McAvoy, and J. Barrows

The effects of differential post-exposure illumination on the decay of a movement after-effect 209
By Irwin M. Spigel

Alternation learning and mental pathology: Test procedures and findings . 211
By Milton B. Jensen

The effect of traumatic experiences on the growth and behavior of the rat . 227
By Charles R. Galbrecht, Roscoe A. Dykman, and John E. Peters

A critical note on Barber's case-study on "Subject J" 253
By Ronald E. Shor and Joel Schatz

The perception of a depth interval with binocular disparity cues . . . 257
By Walter C. Gogel

The determination of perceived movement direction 271
By Martin Krampen and Hans H. Toch

Relationships between extraversion, neuroticism, and ascendance . . . 279
By A. Bronzaft, R. Hayes, L. Welch, and M. Koltuv

A further study of weak-light reinforcement and response facilitation . . 287
By Franklin A. Nash, Jr., and William F. Crowder

Similarity as a factor in recall: I. The effect of color 291
By Dorothy Irene Marquart

An evaluation of the magnitude estimation technique 303
By Glenn R. Hawkes

Phenomenal distance in scenes with independent manipulation of major and minor items 315
By S. Howard Bartley and Doris C. DeHardt

Certain chromatic and brightness changes associated with rate of intermittency of photo stimulation 323
By S. Howard Bartley and Thomas M. Nelson

Some relationships among personal adjustment self-ratings, self-scores, and assigned "average" scores 333
By George F. J. Lehner

A study of incidental learning 339
By M. E. Tresselt and M. S. Mayzner

Switching behavior in a problem solving task 349
By M. E. Tresselt and M. S. Mayzner

The significance of dysphasia for intelligence and adaptive abilities . . 355
By Ralph M. Reitan

A multidisciplinary approach to the diagnosis and therapy of the nonrehabilitated alcoholic personality 377
By Cyril C. O'Brien

Marriage and premarital conception 383
By Horace Gray

Nonlinearly related measurement scales 399
By Eugene S. Edgington

Printed as a separate and in *The Journal of Psychology*, 1960, **50**, 3-14.

METHODS FOR COMPARING THE LEGIBILITY OF PRINTED NUMERALS*

Dunlap and Associates, Inc.

IVAN FLORES[1]

A. PROBLEM

The purpose of this paper is to describe methods used to compare the legibility of printed numerals. The need for this arose during the development of type to be read automatically by machine. Although the basic form of the characters is similar to that of arabic numerals, they are so stylized as to be readable by automatic machine methods. The differences that arose from this kind of stylization are quite distinct from those which arise from consideration of aesthetic appeal. The criterion for stylization here is the machine's ability to read the characters. Since these differences are so unique, it is difficult to make a correct snap judgment as to the legibility of the type. It is desirable to be able to rate relatively the legibility of the type necessitated by stylization for machine reading.

In this system, the occasions when the characters will be read by human beings are infrequent. One instance is for proof reading of the information at the original printing. The second instance arises when lists of documents do not total correctly. It is then necessary for a human being to go back over the documents and check each individually to find the error. Besides a relative rating of the type, it is desirable to find the effect of stylization in this application.

The experiment is divided into two parts. A highly sensitive test for relatively rating the styles of characters, using the tachistoscope, is first applied. The second test is a practical test simulating the actual environment for which the type is intended.

B. SENSITIVE TEST

In choosing a test to compare various styles of type, it is necessary to find a test which will rate even the best type comparatively. By changing various

---

* Received in the Editorial Office on February 25, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to acknowledge the large job of organization and presentation of the experimental work which was completely handled by Saul Teichman. Mechanical layout and construction of the apparatus was efficiently and effectively carried out by Henry Finney despite the existence of only meager sketches of the proposed equipment.

factors influencing the tachistoscopic presentation of numerals, it is possible to force errors on even the most legible type. This is the technique that was used to compare type designed for the greatest legibility rather than for machine readability. At the same time, the test must have some discrimination in the region of poor type. The tachistoscope offers a great latitude because there are several parameters which can be affected so as to change its characteristics.

1. *Equipment*

The method of presentation—the appearance of the image for the subject—is controlled for this tachistoscope by switching on the illumination.

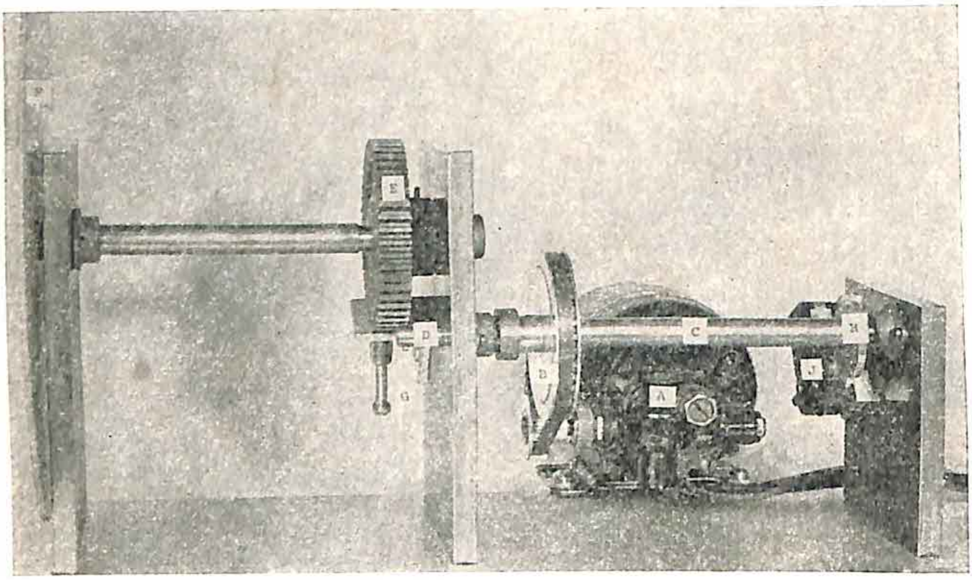

FIGURE 1
SIDE VIEW, TACHISTOSCOPE MACHINERY

A picture of the apparatus appears in Figure 1. The entire mechanism is driven by a speed reduction motor, *A*, rotating at 75 RPM. The speed is further reduced by a reduction pulley arrangement, *B*. Thus shaft *C* rotates at 25 RPM. A single tooth has been set into the shaft at *D*. The tooth at *D* will have no effect upon the gear, *E*, except for a small portion of the revolution of the shaft, *C*. *D* and *E* are so spaced that when *D* does come in contact with the gear, *E*, it drives it for approximately the distance of one of the teeth of the gear, *E*. To make sure of accurate positioning of the disk, *F*, the ratchet detent, *G*, applies pressure to the gear, *E*. Affixed to the shaft, *C*, is a cam, *H*.

As the shaft rotates, the cam, *H*, causes the microswitch, *J*, to operate for a fixed period which occurs some time after the positioning of the disk, *F*.

The microswitch, *J*, controls the voltage applied to the fluorescent tube used for illumination. This is seen in Figure 2 at *K*. A standard 15-watt white fluorescent tube was used. The filaments were kept heated by a six-volt filament supply. An ionizing voltage was applied across the length of the tube. This adjustable DC voltage was set for a proper illumination level. It had to be low enough so that no after-glow remained on the tube after the microswitch broke the circuit. The length of time for which the tube was illuminated was controlled by a mechanical setting in the microswitch placement.

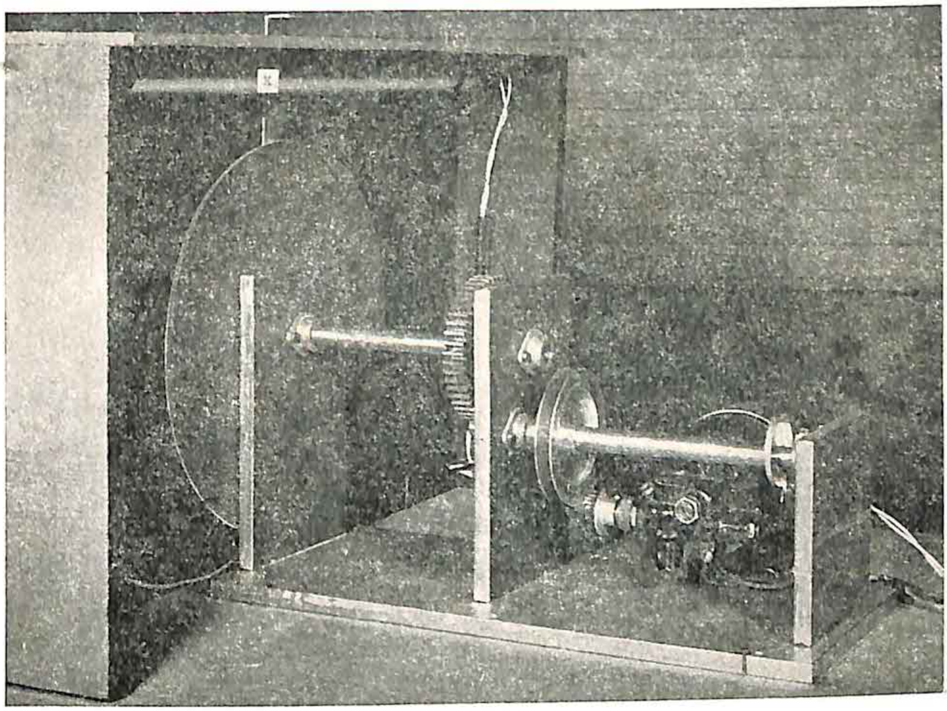

FIGURE 2
ANGLE VIEW, TACHISTOSCOPE MACHINERY

The disk rotated behind a square opening which allowed only one numeral to be visible to the subject, as in Figure 3. The presentation appeared behind the rear wall of an enclosed box, as in Figure 4. The interior of the box was painted flat black so that no reflections would occur from the sides of the box. It was found necessary to allow just enough light leakage during the inter-presentation period, so that the position of the presentation square could be

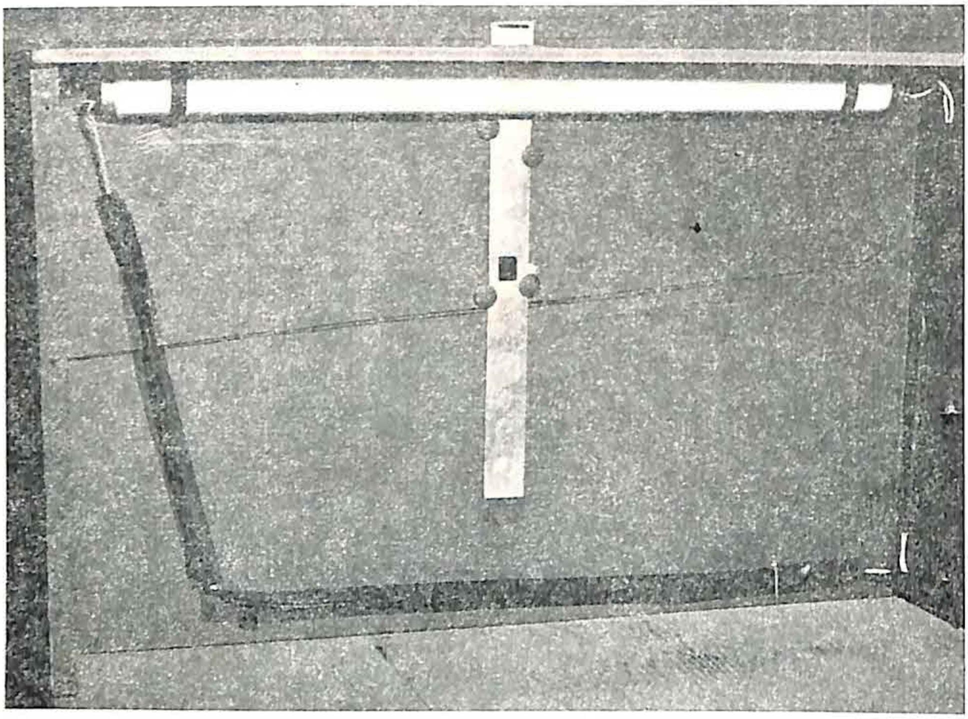

FIGURE 3
REAR VIEW, TACHISTOSCOPE BOX EXPOSED

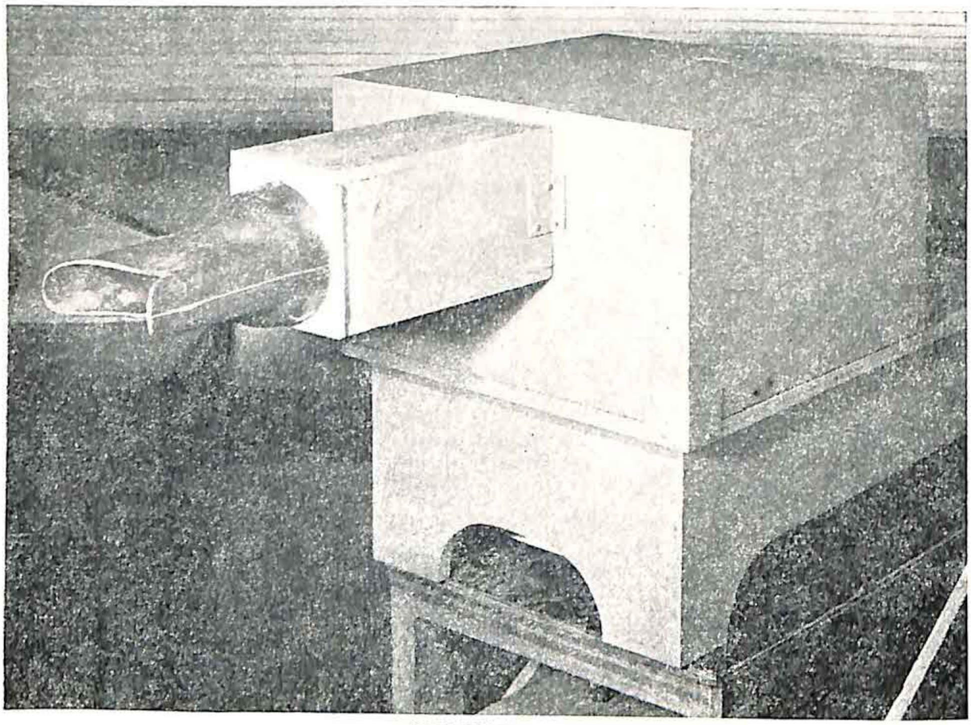

FIGURE 4
FRONT VIEW, TACHISTOSCOPE BOX

observed by the subject (but not the numeral which appeared there). Otherwise the phi-effect would occur and the subject would lose his frame of reference. The next presentation might then startle him because of its apparently different position and cause him to make an error.

Only one style of type appears on each disk as in Figure 5. If more than one style of type were placed on the disk, two rings would have to be used and it would be difficult to illuminate uniformly two such rings. Each disk has 40 characters placed randomly around the presentation area but in such a manner that each character appears exactly four times.

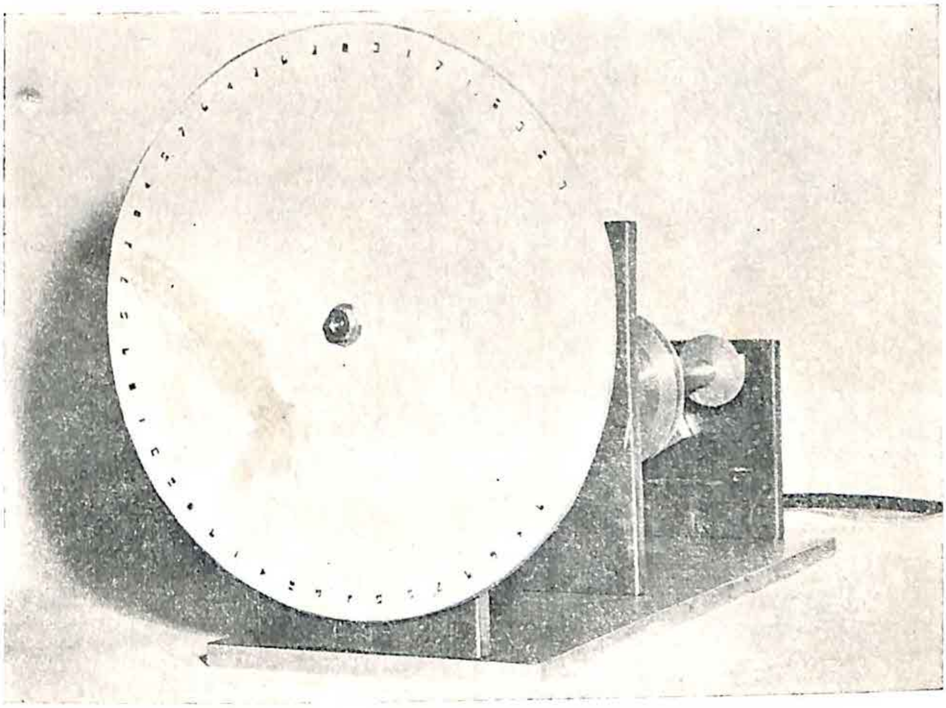

FIGURE 5
Front View, Tachistoscope Presentation Wheel

2. *Procedure*

Each subject was informed of the purpose of the experiment. A typed sheet of instructions was read to each subject. It was emphasized here that it was necessary for the subject to guess if he was not sure of the numeral presented. The subject was then shown a sample of the type to be used in this particular run. The type was arranged in numerical order from zero to nine.

The subject was shown the equipment and he then looked at an illuminated number. He was given a sample run to familiarize him with the procedure. The complete run of 40 characters of one style was given to the subject. Each character was illuminated for approximately 10 milliseconds. During the two and one-half seconds between presentations the subject announced the number he had just seen and the experimenter recorded it. The fonts used are shown in Figure 6.

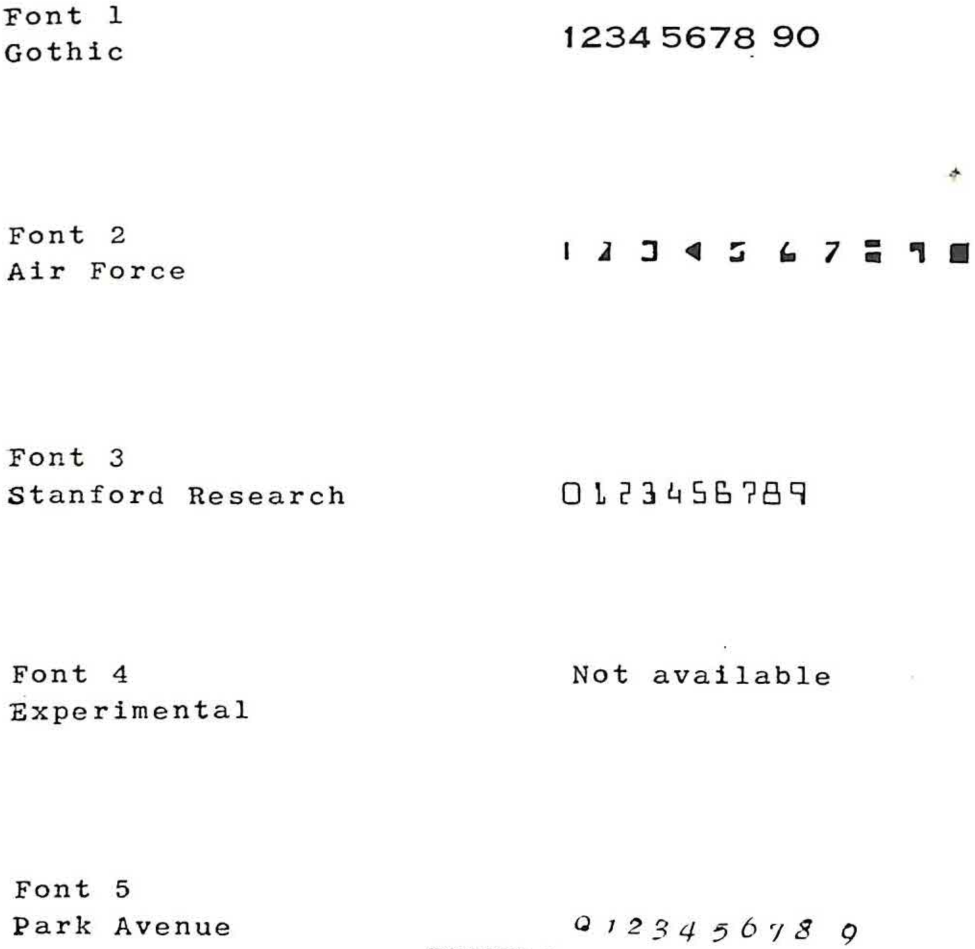

FIGURE 6
TYPE FONTS USED IN EXPERIMENTS

3. *Results*

The number of errors for each run and each subject was recorded and an analysis of variance was run for the entire experiment. The *F*-test was applied to the data. It showed a significant difference among the treatments. A t-test was then used to derive significant differences between means at various presentation points. These results are presented in Tables 1.1, 1.2, and 1.3.

The differences among the treatments vary significantly except for the difference between Font 3 and Font 4.

C. Practical Test

The object of this portion of the experiment was to simulate the situation in which the documents printed in the type in question would be read by the human operator. An example of such a situation occurs when a clerk in a bank has to read and compare a number of documents. He must be able to recognize the number that appears on each one within the time allotted for his handling of them. This is the situation to be simulated.

1. *Equipment*

For each font of type, 25 documents were prepared to simulate checks. The stock for each check was exactly the same. A number was placed in the same area on each check. Each number consisted of three to six digits. The number of digits on a check was obtained by entering a table of random numbers. The first digit of the random number if six or less and three or more would indicate the number of digits on the check in question. The remaining digits of the random number would be used for the digits on the check up to the number indicated.

The number of digits on the check was chosen to be between three and six because in banking practice checks are most often in amounts using that number of digits.

The area chosen on the check was that customarily used for the amount in banking practice.

2. *Procedure*

Before each run, the subject was given a sheet with the numbers of the font listed in ascending order. This enabled the subject to familiarize himself to some extent with the font to be used. The subject was given standard instructions, being told to look at each check, read off the number he saw there, and then turn it over to expose the next number. He was then given a warm-up sample batch of five checks which he read off. A full batch of 20 checks, all of one font, was given to the subject and at the word "go" he would read off each check as instructed. The experimenter would then follow this process. He would follow the numbers being read on a list. If an error in reading was made which was not corrected, he would note it on his list.

TABLE 1.1
ANALYSIS OF VARIANCE, FIRST EXPERIMENT

Variation	Number in each group	Sum of squares	Mean square
Fonts	$N = 27$	$Q_F = \sum_1^5 (\sum_1^{27} E_{ij})^2 = 631,300$	$Q_F/N = 23,300$
Subjects	$P = 5$	$Q_p = \sum_1^{27} (\sum_1^5 E_{ij})^2 = 94,345$	$Q_p/p = 18,302$
Total	1	$Q_T = \sum \sum E_{ij}^2$	$= 28,302$
Error sum	$NP = 135$	$Q_E = (\sum \sum E_{ij})^2 = 2,190,400; \quad \frac{Q_E}{NT}$	$= S = 16,230$

TABLE 1.2
F-TEST, FIRST EXPERIMENT

$Q_T - S$ 12,072

$\frac{Q_F}{N} - S$ 7,080

$\frac{Q_p}{T} - S$ 2,630

9,710

9,710

$Q = 2,362$

$$F = \frac{\frac{Q_F}{N} - S}{(T-1)} \Bigg/ \frac{Q}{(N-1)\ (T-1)}$$

$$= \frac{7080}{4} \Bigg/ \frac{2362}{4.26}$$

$$= 78$$

$F_{1\%}\ (100,4) = 3.51 << 78$ Therefore results highly significant.

TABLE 1.3
t-TEST, FIRST EXPERIMENT

Average errors for each font		Confidence limit for the difference of two means, Δ_p using *t*-test
6.97	Font 1	$\Delta_p = \pm t_p \sigma T \sqrt{\frac{N_i + N_j}{N_i N_j}}$
2.19	Font 2	$\Delta_{1\%} = 6.72$
11.00	Font 3	$\Delta_{5\%} = 5.08$
10.58	Font 4	$\Delta_{10\%} = 4.27$
24.10	Font 5	

Only Font 3 and Font 4 are significantly similar.

One run of 20 checks was made for each of the five fonts in question. These fonts are shown in Figure 6.

3. *Results*

The time for each run was recorded and an analysis of variance was run for the entire experiment. The *F*-test was applied to the data. It showed a significant difference among the treatments. A *t*-test was then used to derive significant differences between the means at various percentage points. These results are tabulated in Tables 2.1, 2.2, and 2.3.

The differences among the treatments were shown by the *t*-test to be due mainly to the effect of the Air Force font. The largest difference among the other four was shown to be due to chance at least 10 per cent of the time; other differences among these other four would be due to chance most of the time. These results are shown in Table 2.3.

D. DISCUSSION

The results obtained from these two experiments seemed at first glance to be completely different and antagonistic. In order to find an explanation for this difference it is necessary to analyze what occurs with the subject in each experiment.

1. *Tachistoscopic Presentation*

In the first experiment the subject gazes at a darkened area. The image of a numeral is illuminated for an extremely short period of time. The image of this numeral appears on the retina of the subject's eye. The nerve endings in the light and dark area are stimulated. Those in the light area are stimulated for such a short period of time that they respond on a probability basis. The image on the retina is transmitted to the recognition area of the brain again on a probability basis.

Now fonts which require distinction of a large amount of detail for recognition will be discriminated against. Because of the probability mode of detection some of this fringe detail is likely to be lost. On the other hand, fonts which depend only upon large components for their identification will be favored. This is because, on a probability basis, sufficient information will usually be transmitted. The comparatively large time (2.4 sec.) between presentations is sufficient for full comparison between the received image on the retina with the full repertoire of recalled comparison images.

2. *Check Reading*

In this case the time for each operation is voluntarily assigned by the subject. During each check reading operation, several processes take place. The

TABLE 2.1
Analysis of Variance, Second Experiment

Variation	Number in group	Sum of squares	Mean square
Fonts	$N = 27$	$Q_F = \sum_1^5 (\sum_1^{27} T_{ij})^2 = 5{,}990{,}000$	$Q_F/N = 222{,}000$
Subjects	$P = 5$	$Q_p = \sum_1^{27} (\sum_1^5 T_{ij})^2 = 1{,}090{,}000$	$Q_p/p = 218{,}000$
Total	1	$Q_T = \sum \sum T_{ij}^2$	$= 230{,}500$
Error sum	$NP = 135$	$Q_E = (\sum \sum T_{ij})^2 = 28{,}070{,}000$; $S = Q_E/135$	$= 212{,}000$

TABLE 2.2
F-Test, Second Experiment

$Q_T - S = 18{,}500$	$F = \dfrac{\left(\dfrac{Q_F}{N} - S\right) \Big/ (P-1)}{\dfrac{Q}{(N-1)\ (P-1)}}$
$Q_F/N - S = 10{,}000$ $Q_p/p - S = 6{,}000$ $16{,}000$	$F = 104$
$16{,}000$ $Q = 2{,}500$	
$F_{1\%}\ (100,4) = 3.51 < < 104$	Differences among times are therefore highly significant.

TABLE 2.3
t-Test, Second Experiment

Average time for each font		Confidence limit for the difference of two means, Δ_p, using *t*-test
33.6	Font 1	$\Delta_p = \pm t_p \sigma_T \sqrt{\dfrac{N_i + N_j}{N_i\ N_j}}$
56.2	Font 2	$\Delta_{1\%} = 8.32$
38.1	Font 3	$\Delta_{5\%} = 6.30$
35.9	Font 4	$\Delta_{10\%} = 5.32$
35.0	Font 5	

The differences among the fonts numbered 1, 3, 4, 5 could be explained by chance more than 10 per cent of the time.

The difference between Font 2 (Air Force) and the others could be explained by chance less than 1 per cent of the time.

subject must scan an average of 4.5 numerals. He must then make a comparison between each numeral he has scanned and each of the possible 10 reference numerals. He must then make a choice assignment for each of the 4.5 possible numerals. He must then cause his speech effectors to clearly announce the choices he has made. He must then turn over the check he was just examining to reveal the next check to be read. Notice here that all these processes are going on at the same time. Also a correction and verifying process is occurring. The subject often notes that the identification of one of the numerals has not been made or spoken correctly and therefore must be corrected. This may be due to an error in initial identification or an error in calling forth the proper spoken word.

The following calculations can be made: the average time for presentation is found by dividing the average time for all the subjects by the number of presentations. For most fonts this calculation yields $36/20 = 1.8$ sec. The average time per character is therefore $1.8/4.5 = 0.4$ sec.

The amount of time available to the subject for the identification of each numeral was less than 0.4 sec. in the case of all fonts except the Air Force. This comparison time was much shorter than the 2.4 sec. available in the tachistoscopic presentation. This time was enough to identify the numerals which were sufficiently similar to the arabic numerals encountered throughout one's daily experience. This is the time required to examine the pattern on the check and match it with the patterns conjured up from the repertoire of reference patterns. When the patterns to be identified depend upon a newly acquired set of criteria, the time required to make the comparisons will be necessarily longer than that required for the familiar pattern.

Thus, for the Air Force font, the time per presentation is $56/20 = 2.8$ sec. and the time per character is $2.8/45 = 0.62$ sec.

What has been determined by the second experiment is therefore the minimum time required for identification of the characters of a given font with very little learning present but under good presentation conditions. The first experiment determines the distinguishability of the characters of a given font under marginal presentation conditions but with adequate allowance time for the comparison operation.

E. Conclusions

The results of the first experiment indicate that under conditions of marginal illumination and observation period but with sufficient comparison time for a given learning situation, the distinguishability of a type font depends mainly on large distinctions, large pattern differences among the character patterns. The second experiment shows that given limited judgment time but adequate conditions for presentation, identification depends to a great extent upon the similarity of the type to that dealt with in daily experience and consequently upon the amount of learning present.

It is conjectured that if learning were continued with the Air Force font to the point where the recognition time for this font was comparable to the other fonts (if this were possible) that the Air Force font would then rate as well in the second experiment as the other fonts. It is possible that the Air Force font would require even less time for identification because the criteria for distinction are fewer and are more easily arrived at. This conjecture could be easily verified by trying to carry the learning of the Air Force font to the saturation point and then performing the second experiment.

When the situation calls for numerals to be identified with a minimum of learning and under favorable presentation conditions, a font with close resemblance to familiar arabic numerals is recommended.

Where immediate identification of numerals under marginal lighting and examination circumstances is necessary, it appears that the Air Force font is superior.

REFERENCES

1. BURT, C., *et al.* Psychological study of typography. *Brit. J. Stat. Psychol.*, 1955, **8**, 29-56.
2. CROOK, M., *et al.* The design of digits. USAF WADC Tech. Rep., 1954, No. 54.
3. LANSDELL, H. Effects of form on the legibility of numbers. *Cand. J. Psychol.*, 1954, 77-79.
4. LUKIESH, M., *et al.* Criteria of readability. *J. Exper. Psychol.*, 1940, **27**, 256-270.
5. MELCHER, D. Psychologists ask: What constitutes "legibility" in type? *Pub. Week.*, 1956, **170**, No. 6.
6. PATERSON, D. G., & TINKER, M. A. Studies of typographic factors influencing speed of reading. *J. Appl. Psychol.*, 1929, **13**, 120-130, 205-219.

Dunlap and Associates, Inc.
429 Atlantic Avenue
Stamford, Connecticut

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 15-38.

PROJECTIVE METHODS RECOMMENDED FOR USE WITH THE BLIND*

Jacksonville Child Guidance and Speech Correction Clinic and Richmond Social Service Bureau

DELL LEBO AND ROSELYN SHERMAN BRUCE

A. INTRODUCTION

The number of projective tests actually used or recommended for use with blind individuals is difficult to determine. It has been described as "deficient" (Braverman & Chevigny, 1955, p. 2), "severely limited" (Lebo & Harrigan, 1957, p. 339), and "scant" (Morris, 1950, p. 114). Dean, in a study "concerned [in part] with the problems of just what tests to use in evaluating adjustment to blindness" (1957, p. 171) used only two projective tests; while the originator of a new tactile projective device for the blind remarked cryptically that no other projective technique than his was available for the testing of blind subjects.[1]

A more mysterious sequence of test recommendations and references has its start in a section on the psychological examination of blind persons in a recent text on testing procedures (Anastasi, 1957). That section served to recommend a manual on testing the blind psychologically (Bauman & Hayes, 1950) which devoted one sentence to projective techniques. The manual, in turn, cited a chapter in a book (Morris, 1950) based on a 1947 conference.

This chapter seems to have been the only attempt actually to list and describe suitable tests. The review cited only nine projective tests, one of which apparently has not yet appeared in published form. Since the enumeration appeared the projective testing of the blind has, according to two observers, "undergone a noticeable development" (Raskin & Weller, 1953, p. 18). The present writers now agree. New projective tests have been reported in the literature, several of the older tests actually have been used with the blind, and normative and validating studies on blind and sighted subjects have appeared.

---

* Received in the Editorial Office of another journal on April 6, 1959, and now given immediate publication in this journal under financial assistance from the American Foundation for the Blind, Inc. Copyright by The Journal Press, Provincetown, Massachusetts.

[1] E. F. Kerman, personal communication, May 26, 1958. Since the present review was accepted for publication this statement has appeared in an article not listed among the references (Kerman, E. F. Cypress knees and the blind. *J. Proj. Tech.*, 1959, **23**, 49-56.). In that article the statement has been amplified and apparently restricted to forms that can be explored tactually since "earlier efforts were not carried very far . . ." (p. 56).

Previous work on techniques deemed suitable for use with the blind has been available in publications largely intended for counselors of the blind. The general literature of psychology contains only scattered references to individual tests and experiments. Since any clinical psychologist may be called upon to test a partially sighted or blind person, and many experimental psychologists may be interested in utilizing such people as research subjects, the present writers believe that a current evaluation of the field is needed.

In addition to describing projective methods suitable for the psychological examination of the blind or partially sighted, the present paper also evaluates the field, indicating seeming weaknesses in an abandoned proliferation of stimuli, and suggesting what seems to be a much needed long range program of research. This evaluation schema might well be useful throughout the realm of projective psychology.

B. AUDITORY PROJECTIVE TESTS: STIMULI READ

Auditory techniques in projective testing are of especial interest because they may also be used by any person who can hear, without regard to visual defects. So they have the possibility of being important additional, but little known, tools for clinical or research work on sighted individuals as well as blind.

1. *Insight Test*

One such method, the Insight Test (Sargent, 1953) originally developed for use with sighted persons, has aroused interest as a useful projective technique for the blind. First described in a monograph published in 1944 (Sargent, 1944) it has been regularly employed as one of a battery of tests used in evaluating the adjustment potential of blind persons undergoing rehabilitation training (Sargent, 1956).

The Insight Test consists of a series of briefly described situations somewhat resembling Murray's (1943) verbal descriptions of the Thematic Apperception Test (TAT) pictures. At present the Insight Test consists of four forms, each with 15 psychological situations described. While as yet there have been no specific problems introduced into the test for blind subjects (Sargent, 1953, p. 227), the following statement is an example of the type of verbal picture suggested for such use: "A woman who has been blinded in an accident receives a letter from her daughter asking her to come to live in her home."

Subjects are instructed to analyze what people do and feel under the circumstances set forth. They are also told that there are no right or wrong answers but that they should show insight into the actions of the characters.

Sargent (1953) realized that an unlimited number of problems could be concocted for blind individuals. However, she cautioned against using more than a few in an Insight Test. She advocated, instead, a majority of regular questions so that the examiner's experience with the original material could be put to advantage to judge alterations in form and content when personal problems were approached. She also surmised "that questions about deafness, for example, or other handicaps, might be used in forms for blind persons . . . since experience has shown that in some instances attitudes embedded in highly sensitive areas appear less disguised when the projective figure is somewhat remote" (Sargent, 1953, p. 228).

Examples of the use of 10 items of the Insight Test on two blind persons are available with full protocols and analyses (Sargent, 1953). Methods of analysis, as is so often the case in projective psychology, range from content analysis to a "cumbersome" (Sargent, 1953, p. ix) formal procedure. The latter scoring system involves differentiating between expressions serving to control, defend, or delay unmodulated emotions. The subjects' phrases are categorized with regard to the outcome of aroused affect and expressed mood. Relationships between some expressions said to represent affect and defense or discharge or control may be computed. In addition, Sargent expects the "clinical research worker" to "take creative liberties with . . . the method of analysis . . ." (1953, p. ix). One such worker has developed a simpler system (Fassett, 1948).

Another worker, who collected 75 or more records from blind subjects, reported that the Insight Test was "the most promising method presently available [1947] for research with the blind" (Morris, 1950, p. 125). Dean, however, comparing Sargent's findings with protocols from 54 blind persons, concluded "that the Insight patterning of answers suggests a cautious use of [Sargent's] norms with the blind" (1957, p. 177). Sargent herself, in a study of 27 blind persons, called attention to "the danger of relying, at this stage of knowledge," on quantitative findings (1956, p. 441).

2. *Verbal TAT (VTAT)*

Like most clinicians, Sargent has found that the scoring methods used in conjunction with one projective device may profitably be applied to another. Methods used in conjunction with the TAT, for example, were said to be "especially applicable" (Sargent, 1953, p. 21) to the Insight Test. By the same token, Insight Test scoring method has been applied to TAT protocols (Sargent, 1953).

As may be inferred, the two methods have much in common. Indeed, a verbal form of the TAT has been in use with the blind since approximately 1952.[2] This use of a VTAT is by Sammie K. Rankin, and is based upon her modifications of the TAT descriptions in Murray's *Manual* (1943). An examination of Rankin's VTAT[3] and descriptions given by Murray reveals many interesting changes. In view of the use of Murray's descriptions in preliminary investigations of another development of a VTAT for the blind, to be discussed in turn, two of Murray's TAT descriptions and Rankin's VTAT stimuli are presented for comparison. First, two similar descriptions based upon Card 12BG are compared, next two dissimilar verbal depictions, based upon Card 19, are contrasted. To make comparison and identification easier Murray's (M) description is given first followed by Rankin's (R) in each case.

> (*M*) 12BG. A row boat is drawn up on the bank of a woodland stream. There are no human figures in the picture.
>
> (*R*) 12BG. A row boat has been drawn up on the wooded bank of a stream. There are no people in this scene.
>
> (*M*) 19. A weird picture of cloud formations overhanging a snow-covered cabin in the country.
>
> (*R*) 19. This is an abstract picture. Everything is blurred in outline as if with snow, and wind is indicated by jagged projections on several blurs; these projections stream in the same direction. There is a mound of white with two, lighted squarish spots which might be windows or doors. There are two more lighted squares high above the first two. There are blobs and splotches of gray at the top of the picture.

Rankin uses standard instructions for her VTAT, which she has found to be "an excellent projective technique." She prefers a content analysis approach although she occasionally has used that of formal scoring.[3] Further, she has "observed no significant difference between the stories of blind persons and sighted persons even when the responses from sighted persons have been given to the picture instead of the verbalization."[4] These observations, however, are based largely on impressions or inspection rather than on statistical evaluation.

Independently of Rankin, though at a later date, Lebo (1955) noted that disliked TAT pictures were usually the same cards whose descriptions in Murray's *Manual* (1943) contained disquieting elements. This observation suggested the possibility that the original descriptions of the TAT cards contained approximately the same stimulus value as did the cards themselves.

---

[2] Sammie K. Rankin, personal communication, April 9, 1958.

[3] Sammie K. Rankin, personal communication, May 1, 1958. Since the present review was accepted for publication a manuscript listing Rankin's entire VTAT has been scheduled for appearance in this Journal (Lebo, D. The development and employment of VTAT's or pictureless TAT's. *J. Gen. Psychol.*, in press.).

[4] Sammie K. Rankin, personal communication, June 12, 1958.

Lebo and Harrigan (1957) later compared TAT and VTAT presentations of stimuli. It was hypothesized that if card descriptions were read to subjects the same responses might be elicited as from the usual TAT administration. To compare the value of such stories to traditional TAT protocols they used various objective measures including word count, idea count, emotional tone of stories and outcomes, response level, dynamic content, perceptual range, common themes, and normative data. All of these devices had been successfully employed in the traditional TAT literature.

It was found that the verbal descriptions of TAT pictures evoked responses comparable to those of the TAT itself. The substitution of verbal descriptions for visual plates was apparently justified, at least for experimental purposes.

A later study by Lebo and Sherry (1959) contrasted a VTAT read to subjects with a VTAT read by subjects. Again, Murray's unaltered descriptions were used. Differences between these two methods of VTAT presentation were largely nonsignificant according to the objective methods of comparison listed previously. It was emphasized, however, that the TAT appeared to possess stimulus properties not evident in Murray's descriptions. Rankin's VTAT apparently was developed to close such a gap.

While Lebo and his coworkers have not yet used a VTAT on blind subjects they have been concerned with such usage ultimately. It was felt, however, that individual comparisons of protocol content to TAT plates and descriptions were necessary before developing a substitute VTAT. The more the VTAT parallels the TAT the easier it should be to use TAT norms in VTAT interpretation. Research along these lines is underway and a preliminary scoring manual (Lebo, 1959a) with the presentation and analysis of two 20-item VTAT's has been prepared. In the meantime, material of research interest and clinical value in general has been presented (Lebo, 1959b).

3. *Sentence Completion*

Like the TAT, the Sentence Completion test has been used extensively with sighted subjects before being applied to the blind. The Incomplete Sentence Blank (ISB) has been recommended for use (Bauman & Hayes, 1950; Bean, 1957; Morris, 1950) and actually used (Dean, 1957) with the blind. The ISB is a semi-structured technique which confronts the blind patient with clearly spoken (word or phrase) sentence stems. He is instructed to complete these stems in a manner expressive of his real feelings. Some extremely ambiguous items such as "I . . . . ," and "Sometimes . . . . ," are included.

Originally devised as a military screening test (Rotter & Willerman, 1947), the ISB still consists of 40 items (Rotter, 1950). As with the Insight Test, experienced psychologists are urged to devise special items for their purposes. Thus, Rotter and Rafferty (1950) have changed several items to develop an ISB College Form. By means of a few additional changes this form can be converted into one for nonschool adults or high school students. Additional changes would seem to make such forms easily adaptable to blind persons of various age and cultural groups.

The entire test protocol can be obtained in about 20 minutes on a single sheet of paper. It can be scored by classifying completions into categories such as conflict responses, positive or healthy responses, and neutral responses. The latter are completions not falling into either of the previous categories. Humorous and avoidance responses such as "Most girls *are females*" are regarded as neutral statements. Numerical values are assigned to indicate the nature of the response group and the strength of the statement. Conflict responses are given positive values while positive statements receive negative scores. Thus, the higher the score the greater the maladjustment. Scoring methods and examples are available in the *Manual* (Rotter & Rafferty, 1950).

Rotter and Rafferty (1950) have reported a cutting score which identifies 75-80 per cent of maladjusted subjects. Dean (1957) found that 54 blind persons scored higher than Rotter and Rafferty's norms for a well adjusted group and for college freshmen but lower than maladjusted groups. Although the chief recommendations of the ISB are its objectivity and its lack of dependence upon unusual clinical skill—relatively untrained persons can score it reliably—Dean concluded that, the "Rotter ISB . . . value with the blind probably resides in qualitative rather than quantitative evaluations" (1957, p. 176).

Another form of sentence completion test proposed for use with blind persons (Morris, 1950) is that of Stein (1947). This test consists of two parts, each of which contains 50 complete sentences. A variety of areas important in the evaluation of personality are covered. They include: family, past experience, drives, inner states, goals, cathexes, energy level, time perspective, reactions to others, and reactions to frustration. Sentence stems are of two types, i.e., references to third parties ("Mike's fondest ambition . . . .") and to personal items ("I try hard . . . ."). Responses are analyzed by underlining significant completions. Hence, the interpretive skill of the clinician is emphasized strongly. The test has, understandably, been characterized as "descriptive but not evaluative" (Symonds, 1947, p. 321).

McAndrew (1950) used a sentence completion test with 25 blind, 25 deaf, and 25 normal children matched for age and intelligence. He found the blind group to be most rigid of the three.

An interesting variation in the presentation of sentence completion material is a Braille adaptation (Brieland, 1950). This version is discussed more fully in the section of the present paper on three dimensional tests.

4. *Word Association*

The auditory projective tests treated thus far all have the use of unambiguous words in common. Starting with techniques employing complete statements the survey has progressed through briefer stimuli to the use of single words as stimuli.

Word association is one of the oldest formal free association techniques. It has been recommended for use with the blind (Bean, 1957; Morris, 1950) but a search of the literature has been unsuccessful in providing reports of such use. One of the first to use this method was Jung (1918). His list of 100 words, chosen to represent areas of emotional disturbance, is still among those frequently cited. A list frequently employed is that of Kent and Rosanoff (Rosanoff, 1938). Disturbed associations to their 100 neutral words are detected by unusual content.

While word association tests are easily administered, normative responses have varied so with culture, region, intelligence, and time that they have not proven generally representative. The most recent frequency tables are those of Tresselt and Leeds (1953a, 1953b). For particular research problems it would seem to be advisable to prepare special words. Consequently, norms for such words will also have to be established.

C. Auditory Projective Tests: Stimuli Recorded

Interest in vague auditory stimuli as projective, or imaginative media can be traced to da Vinci (Richter, 1939). He mentioned the confused noises made by bells in which words or names might be imagined. The jangle of bells and the confused or disguised sounds of several auditory projective tests are not too unlike. Auditory projective devices requiring recorded stimuli may be arranged from those making use of clearly audible words, through tests employing inaudible, disguised, or unreal speech and sound effects, to a test employing neither speech nor sound effects.

J-329
3.9.70

1. *Azzageddi Test (Azz Test)*

The peculiar title of this technique comes from Melville's *Mardi and a Voyage Thither* (1922). In that story a devil, named Azzageddi, was blamed for causing the speech of one of the characters to become confused and incoherent. The name of the test, in addition to being appropriate, has another advantage in that the purpose of the test is not inherent in its title. To disguise further its purpose, subjects are informed that it measures "intuitive intelligence" (Davids, 1956, p. 416).

The Azz Test (Davids & Murray, 1955) was designed to measure eight specific emotions or feelings. Believed to be important components of personality structure, these dispositions are: optimism, trust, sociocentricity, pessimism, distrust, anxiety, resentment, and egocentricity. For each disposition eight sentences and six short phrases were recorded. Stimulus material for each of the feelings also contains isolated statements and phrases associated with other dispositions. For example, recorded material for the feeling of optimism contains words associated with other dispositions. Four unrelated statements in the passage are intended to represent trust, distrust, sociocentricity, and egocentricity. Six dispositions are also represented in it by a single word or a short phrase.

The Azz Test recorded stimulus for evaluating optimism is: Beware of sly men, laden with malice, breeders of dirty lies, and smearers of character. Boiling with rage. Opportunities for happiness are unlimited. Disbelieve. It is not what we take up for ourselves, but what we give up for others that truly makes us rich. Trustworthy men. Have confidence in yourself and in your future. Life is a great goblet of glorious possibilities, brimming over with enough delight to make us giddy. Equality for all. Never lose faith in your fellow men. Desire for power. Let neither sorrow nor disappointment bend you down to earth. Peril upon peril. People love themselves above all others (Davids & Murray, 1955, p. 550).

By themselves the aforementioned statements are coherent and unambiguous but when combined the resultant passage is considerably confused. Ambiguity and incoherence is provided by the nature of the content rather than by inaudibility or masking.

Subjects are to remember as many ideas as possible while listening to each of the one-minute passages. A three-minute interval follows each passage during which time the subject is required to write or tell as many ideas, phrases, and statements that he may recall.

Relations between selective auditory memories of the passage above and subjects' personalities were presented by Davids and Murray (1955). Understandably, they feel the procedure may prove useful in studying the mental processes of the blind. Insofar as the literature is concerned this application does not yet seem to have been made. The orginators have, however, compared scores for five negative dispositions, i.e., egocentricity, distrust, pessimism, anxiety, and resentment, on the Azz Test with a tape recorded word association test and a sentence completion test. *Rho's* were .55 and .54 respectively, both significant at the .01 level of confidence.

2. *Auditory Sound Association Technique*

Wilmer and Husni's (1952, 1953) auditory projective test consists of 21 sound series. These evocative noises, including words, were chosen, from 117 originally tested, because they produced greater responses, more varied contents, and also because they were rejected for more significant reasons (Wilmer, 1951). They are available for research purposes on two 12-inch records (Briggs, Gaede, & Wilmer, 1956). Subjects are told to associate meanings to the sounds and to represent aloud the first thing that comes to mind. If subjects wish they may make up stories about the sounds.

Some of the sounds, with the number of their sequence, are:

> 8. Laughter and man and woman conversing. A. "Where did that damn thing come from?" B. "What thing?" A. "Mother said you can't always get what you want in this world, sister." B. "What did you expect?" A. "Where is Dad? Where did Dad go?"
>
> 10. Man and woman arguing (man speaks in anger). Man: "Now what did you do all that for today!" (Mumbling answer). "Huh?" (Mumbling answer). "You know what I'm talking about!" (Mumbling answer).
>
> 1. Train leaving station, crowd, depot noises.
>
> 21. Train bells, switch engine, loud bang, voices.
>
> 15. Groaning, rubbing, kissing (human).
>
> 11. Music and water sounds, ending with splash like fish in water or man drowning.
>
> 18. Drill.

Tests of this nature have been recommended for use with the blind (Allen, 1958) and 21 blind children have been compared with other disabled groups (Wilmer & Husni, 1953). The blind subjects mentioned elements of violence, death, and struggle more frequently than tuberculosis patients, schizophrenic patients, and amputees. The blind children also detected background noises more keenly but tended to reject other components of sounds when their verbalization was fixed on danger elements. Their own handicap was reflected in comments about walking, unseen danger, and warning signals. Individual responses were presented and analyzed.

3. *Auditory Apperception Test*

Stone (1950) has developed an Auditory Apperception Test which has been recommended for use with the blind (Allen, 1958; Stone, 1950). This test also consists of recorded words and sounds of various types. Like the Auditory Sound Association Technique, episodes in dialogue form are also used. Because of the inclusion of discernible phrases the test is considered among those utilizing words.

The set of 10 records, one of the few recorded auditory projective devices available commercially (Stone, 1953), presents three sounds in each of 10 "sound situations." Various combinations of the following types of sound occur: (*a*) dramatic episodes, including a single dialogue such as: "You ready Joe?" "Yeah, I'm ready." "You ready Fella?" "I'm—I'm not sure." "Come on, let's go!" (*b*) non-verbal human sounds (crying, laughing); (*c*) animal sounds (bird calls); (*d*) nature sounds (thunder, fire); and (*e*) mechanical sounds (pounding, bells).

After hearing each group of three sounds, the subject makes up a story, much like the TAT requirements, telling what has caused the sounds, what is happening, and the outcome. A record form, making a TAT-like analysis according to thema, character, and outcome easier, is available.

4. *Auditory Projective Technique (APT)*

Braverman and Chevigny (1955), with the help of the American Foundation for the Blind, have advanced a test believed applicable to persons with severe visual defects. The APT consists of two records with short, spoken sequences and sound effects of unstructured content. The sequences are intended to be auditory depictions of feelings and objects. The first record consists of seven conversations between people of different sexes and ages. Some of the sequences consist of a quarrel, a love scene, and an old couple. Although these conversations have an emotional content they are ambiguous as to word content because they are in a nonsense language. Six subsequent sequences are English language versions of several earlier sounds. The obverse of the second record consists of brief sound effects designed to portray scenes, i.e., storm, escaping criminal, suicide, sea rescue, and snooper.

In addition to being instructed to use imagination, subjects are also to make up a TAT type of story, telling what is going on, what led up to it, and the eventual outcome.

Both the APT and the TAT were administered to 40 patients, excluding disturbed psychotics. Information obtained from the APT was comparable to that secured from the TAT when a need-press type of analysis was made of the protocols. The originators, however, point out that while their test is out of the experimental stage it has not been standardized. APT records are available for research workers (Braverman & Chevigny, 1955).

5. *Tautophone (Verbal Summator)*

The Verbal Summator, introduced by Skinner (1936) is one of the oldest of recorded, free association tests. The Verbal Summator is made up of unfamiliar speech samples. While employing the Summator to investigate hypotheses concerned with the development of speech, Skinner drew an analogy between that device and projective techniques. Shakow and Rosenzweig (1940), stressing projective possibilities, renamed the test as the *Tautophone.* They used only two of Skinner's records consisting of various vowel combinations, each of which was repeated 10 times. (Tautophone is a combination of the Greek *to auto* and *phone,* meaning the same sound.) Some of these sounds are presented below. Long vowels are indicated by capital letters, short vowels by small letters, the vocalization *uh* by ': "I'ah, 'a'A, ahO' ', and Eah ' ' ' (Grings, 1942, p. 541).

Subjects are informed that the recorded voice is unclear and that it will be necessary to listen carefully. As soon as a subject has some inkling of what is being said he is to tell the examiner. Thus, to Eah ' ' ', one subject's interpretation was "my knees are apart" while that of another was "eu as in Hindu" (Shakow & Rosenzweig, 1940).

The test lasts from 20 to 30 minutes. It has been suggested that two persons administer the Tautophone, one to operate the phonograph and record the number of stimulus repetitions preceding a response, the other to record responses and other pertinent information. Shakow and Rosenzweig (1940) have devised a method of scoring which considers such response factors as similarity to stimulus, English language nature, personal references, grammatical structure, i.e., interrogative, imperative, or declarative. They have also developed a series of 12 indices measuring such dispositions as suggestibility, self-reference, subjectivity, and perseveration. Perseveration, for example, may be determined by dividing the number of different words in the responses by the total number of words. The originators of these indices have pointed out similarities between Rorschach scoring and Tautophone evaluation methodology.

Trussell (1939) found the Verbal Summator discriminated between groups of normals and abnormals while Grings (1942) found differences among groups of psychotics. Despite such early usage, no reports were found indicating that such stimuli had as yet been applied to blind subjects. The test is applicable to such persons and has been so recommended (Allen, 1958; Morris, 1950).

6. *Sound Apperception Test (SAT)*

Bean's SAT (1957) seems to be one of the few auditory apperceptive techniques which employs sound effects and no human vocalization. SAT material consists of recorded sounds ranging from ambiguous to semi-structured noises. Ordinary sounds were made ambiguous by distortion, i.e., by slowing their speed, reversing them in sequence, or reducing them in frequency.

The SAT is comprised of two parts, the first consists of 10 unstructured noises and the last of 16 semi-structured sounds. In responding to the first series of sounds, subjects are asked to indicate of what they are reminded by the sound. The second series, however, requires subjects to tell what is happening, what led up to the event, and finally what the outcome might be. Special attention should be given to the thoughts and feelings of any people that might be included in the stories. Once again the influence of the TAT is apparent in the instructions.

One example of a sound series in the TAT like sequence are noises of a diesel train approaching and passing, galloping horses, and then swishing canoe paddles. Some responses to such stimuli describe a pleasant vacation trip, others picture an escaping criminal who jumps off the train, hides as a posse on horseback looks for him, and who finally swims away. These two patterns seem to be basic for the sound sequence. Many variations are said to occur, however.

Preliminary testing comprised approximately 200 persons. Of this number 35 were blind. Such subjects were found to take more interest in the stimuli than did the sighted and to give fuller elaborations. When the variables were quantitatively measured, however, few significant differences were found. Scoring has been standardized and interrater agreement seems stabilized around .73 to .76 (Bean, 1957). Scoring sheets, a manual of instructions, and the recorded stimuli are available.[5]

7. *Musical Reveries Test*

In addition to being relaxing, music can also serve as stimulation to expressive activities, and as a physical or emotional release of conflict (Ashford, 1955). Since the blind are believed to live generally under tension, music has been recommended as an excellent tool for such people (Korhonen, 1956).

A Musical Reverie Test, suggested for such use (Morris, 1950), is that of

---

[5] K. L. Bean, personal communication, June 21, 1958.

Kunze (1938). The test consists of listening to parts of classical compositions while the mind drifts. Subjects are to note their mental images and weave them into a plot or allegory. When the selection is completed an account of the reverie is rendered.

The records were: (*a*) *Symphony No. 4 in F minor,* by Tschaikowsky, which usually suggested fear, agitation, mental or physical struggle. (*b*) *Don Juan,* by Richard Strauss. The last part of the second movement brought to mind love scenes, sentimental or romantic settings. (*c*) *Symphony No. 6 in B minor (Pathetique)* by Tschaikowsky, last movement only. This record brought forth stories of tragedy, despair, and sometimes death. (*d*) *Quintet in G minor,* by Mozart, closing movement. This suggested happy, light-hearted, animated action. (*e*) *Death and Transfiguration,* by Richard Strauss commonly brought to mind feelings of reverence and solemnity, or ideas of regeneration. (*f*) *Afternoon of a Faun,* by Debussy, placid and slow-moving, seemed to facilitate reminiscence and philosophizing. "Other records were occasionally used if one kind of music seemed to be definitely more productive than the others" (Kunze, 1938, p. 551). Kunze reported further that although these selections brought about rather characteristic attitudes and images, their specific form was influenced by the subjects' past experiences.

Despite its early appearance, the present writers have been unable to locate any references to the use of this music test with the blind.

D. TACTILE PROJECTIVE TASKS: THREE DIMENSIONAL TESTS

It is often assumed that the blind acquire superiority in their sense of touch. [Even though Seashore and Ling (1918) reported that they were not more sensitive in sensory discrimination than seeing persons in fundamental capacities.] Hence, it is not surprising to find a variety of three-dimensional projective methods recommended or created for evaluating the personality of the blind. For discursive purposes these tests are arranged on an age-depth continuum. That is to say, the oldest and most truly three-dimensional tests are mentioned first.

1. *Three-Dimensional Apperception Test (3-DAT)*

The 3-DAT aims to elicit the overt expression of subjects' interests, needs, goals, sentiments, and emotions (Twitchell-Allen, 1952). Constructed by Twitchell-Allen (1948), it consists of 28 ambiguous plastic pieces ranging from geometric to organic forms. Her thesis is that manipulation and dramatic use of objects reveals properties of the person not disclosed by verbalization. Indeed, a comparison of 3-DAT and the Rorschach results revealed that more concrete material concerning inner goals, interests, conflicts, and anxieties arose from the former (Twitchell-Allen, 1947).

Administration consists of three parts: (*a*) the psycho-dramatic test in which the subject selects one or more pieces and makes up a story about them. This procedure is repeated several times. (*b*) A naming test in which the subject is asked to name the objects. (*c*) An inquiry in which the subject must indicate the reasons for his assigned names. During this phase subject may also be asked to compose a story centered on each of the names assigned the various objects. Stories and gestures are recorded formally in a summary record. Analysis, however, is left to the examiner's discretion.

As might be suspected, the 3-DAT has been recommended for use with the blind (Bauman & Hayes, 1950; Morris, 1950). In addition, McAndrew (1950) used it with 25 blind, 25 deaf, and 25 normal children. He studied individual personalities with the 3-DAT and a sentence completion test and found that while both tests were applicable to the blind they must be interpreted cautiously.

Test materials and scoring forms are available commercially (Twitchell-Allen, 1948).

2. *Avon Three-Dimensional Apperception Test*

This is one of the few projective techniques developed exclusively by use with the unsighted. Interestingly enough it has been recommended for sighted subjects (Levine, 1950). It consists of figures made in various media, such as plaster or lead, to which subjects free associate. The objects may be named, used as stimuli for stories, or otherwise given meaning. The instructions are simply, "Just associate to the figures, tell everything that comes to your mind while your hands are moving over them." The set is not completed (Morris, 1950).

3. *Kerman Cypress Knee Test (KCK)*

Like the preceding three-dimensional techniques is the KCK (Kerman, 1957). Once again subjects are instructed to describe (or name) each piece and are allowed to select some or all of the forms as characters or properties for a story.

KCK forms are rubber replicas of six cypress knees. Cypress knees are natural above ground growths from cypress tree roots. They are popular souvenirs of Florida because of the myriad ornamental, grotesque, and ambiguous shapes they assume.

Three testing situations make up the administration of the KCK: (*a*) The object choice section in which the subject ranks his preferences for the forms, on a like-dislike basis, and justifies his ranking. (*b*) The part called projective images in which the subject names and tells about the images perceived. (*c*) A story situation during which two plots are required. The first task is a free association to characters already perceived in the KCK. The subject is next asked to assign the roles of mother, father, and child to three of the objects and compose a story involving these assignments.

Analyses, based on Fairbairn's (1954) object-relation theory of personality, consists of determining the symbolic meaning of each form and in assessing subject's attitude towards them. As a result the subject may be categorized as schizoid, depressive, phobic, obsessional, hysterical, paranoid, delinquent, or organic. Further discussion of theoretical aspects and administration as well as analysis of the records of sighted subjects have been presented in detail [Kerman, 1957, in press (*a*), in press (*b*)].

Kerman's initial experience in testing six blind subjects and copies of two full protocols and analyses are available in pre-publication form.[6] He is convinced that the percepts gained by the blind subjects' tactual exploration were practically identical with those formed through visual inspection.

Qualified applicants may secure the test stimuli from the originator (Kerman, 1957).

4. *Bas-Relief Projective Test*

Harris (1947) has constructed a projective test for the blind using bas-relief forms of varying textures, shapes, and structures as stimuli. These forms were based upon a series of chance ink blots, without attempting to imitate the Rorschach (Harris, 1948). Harris' plates were neither all symmetrical nor unstructured. Three of the plates, resembling horses, were initiated to determine the threshold of recognition. The first, clearly structured, resembles a horse picture, the third, almost unstructured, is so disorganized as to resemble only slightly a horse, while the second is in the center of this continuum.

The plates are presented three times. For the first presentation subjects are asked to examine (tactually) the figures and tell what they mean. The second presentation seems like a Rorschach inquiry, i.e., the examiner seeks to understand the nature of the stimulus to which the subject responded, to form or texture, for example. On the final presentation the examiner tries to force implied responses into overt expressions.

The practicability of the Bas-Relief Projective Test has been demonstrated during several experiments with approximately 100 blind and 300 blindfolded, but sighted, subjects. Unfortunately, much of this work seemingly has gone unreported in the literature.

---

[6] E. F. Kerman, personal communication, May 26, 1958. Since the present review was accepted for publication one of these records has been published. See footnote No. 1 for bibliographic information.

5. *Braille Sentence Completion Test*

Brieland (1950) gave a sentence completion test to 250 visually handicapped children who were matched with sighted controls. Thirty items concerning social and home adjustment and family attitudes were presented in Braille or large print. Subjects wrote their answers.

The blind or partially sighted persons showed increased emotionality, excessive anxiety about family relationships, and more dissatisfaction with school achievement.

E. TACTILE PROJECTIVE TASKS: CREATIVE AND EXPRESSIVE

The projective techniques discussed in this section involved something the subject has made, i.e., sculptured, painted, drawn, or written. Since these procedures are more unstructured than any of those mentioned previously, it is not unexpected to find that methods of interpretation are similarly ambiguous and depend to a greater extent upon clinical intuition and acumen than many of the previous tests.

1. *Plastic Material*

In his listing of projective techniques of value in testing the visually handicapped, Morris (1950) called attention to projective techniques frequently employed with the sighted. Quite properly he included clay modeling in the enumeration. However, in saying that, "practically all of them are unsuited for use with the blind, vision being an important factor" (1950, p. 117) Morris seems to exclude plastic material from consideration in revealing the personality dynamics of the blind. Bauman and Hayes (1950), following their one-sentence treatment of projective techniques, cite an article (Bender & Woltman, 1937) on plastic material. Such citation, to the present writers, connotes a recommendation.

Bender and Woltmann (1937), while not working with blind subjects, present much theoretical, explanatory, and interpretative material of value in reflecting personality dynamics. This diagnostic aspect, unfortunately, has not been one of Lowenfeld's (1939) interests in his published studies of a large number of the clay models, drawings, and paintings made by blind and partially sighted children and adults. He has been concerned with aspects of *Creative and Mental Growth* (1955) rather than with the entire personality.

Nonetheless, his "discussion of these factors and his exposition of them . . . is highly stimulating and should be read by all who are interested . . . in projective theory . . ." (Goodenough & Harris, 1950, p. 411).

2. *Finger Painting and Handwriting*

The first medium has been especially recommended (Kadis, 1952) and used (Napoli & Harris, 1948) with the blind because of the plasticity of the material. It allows for freedom in manipulation while the picture itself is a socially acceptable projection of needs. According to Napoli and Harris (1948) the very motion of smearing is relaxing and there is no social structuralization to prevent expression.

Work with 25 blind or partially sighted subjects revealed that their motions were not limited by their handicap but by their emotional set instead. The texture of the pictures was similar to those produced by sighted persons. Composition, balance, and perspective were also well represented (Napoli & Harris, 1948).

It is interesting to note that Wolff (1948, pp. 156-158) in studying the handwriting of the blind, came to much the same conclusion.

The methods of administration and interpretation of finger painting vary according to the examiner. Subjects may be asked to make up a story around their pictorial construction. A simple, objective 12-point scale differentiating between nonpsychotic, manic depressed, schizophrenic, and organic groups has been developed (Lehmann & Risquez, 1953), but has not yet been tested on blind persons.

F. DISCUSSION

Like the majority of projective techniques, those treated in the present paper are largely in an experimental stage and depend in great measure upon clinical skill rather than routine scoring. In several cases studies concerned with validity and reliability have been performed and objective scoring schemes advanced. However, because examiners of the blind are frequently urged to alter tests to suit particular needs, reliability and validity data, unless of particular interest, were not treated in detail. The reliability of one form of a test should be established anew each time any segment of that test is modified. It is regrettable that this precaution is undertaken only rarely, even with one of the most frequently used and frequently altered objective tests (Lebo, Noblin, & Toal, 1957).

This survey of tests considered effective by investigators who have indicated their recommendation in print was concerned with devices stimulating projectivity or creativity (the terms are almost inseparable to the present writers) through senses other than sight. The avenues of perception which remain are hearing, touch, smell, and taste. Little traffic has been reported on the last two avenues. None of what has been published is concerned with the blind.

Several of the auditory projective tests requiring recorded stimuli have names so similar as to be confusing: Auditory Sound Association Technique, Auditory Projective Technique, Auditory Apperception Test, and Sound Apperception Test. The name Auditory Sound Association Technique, used by the present writers, comes from the title of the first publication describing this test (Wilmer, 1951). This designation was not employed consistently in other articles by Wilmer and Husni (1952, 1953). Since the originator has not yet named his device its name may be altered easily. The Sound Apperception Test (Bean, 1957) also seems to be a tentative title. Names of these two tests, then, may be altered without difficulty.

To prevent increasing confusion, as more auditory projective devices appear, the present writers suggest a uniform policy in regard to designation. Since the tests above have many similarities and equal evanescence a subspecies-species designation would seem to be applicable. The names of the senior investigator should precede a more general, abbreviated indication of the *auditory projective* nature of the *test*. Thus, the Auditory Sound Association Test becomes Wilmer's APT, and the Sound Apperception Test becomes Bean's APT. This usage is consonant with that employed to differentiate Rankin's VTAT stimuli from Murray's descriptions and may be extended to embrace the established 3-DAT designation. Unusual identifications such as the Azzageddi Test probably will cause no difficulties, unless the test becomes popular.

The 1947 review of the literature (Morris, 1950) did not encounter this difficulty with names. It described only the Tautophone and the Musical Reverie Test. A variety of APT's has been introduced since then. The literature reporting these new arrivals is much alike. They usually begin by reviewing publications on previous APT's. Immediately complaints about the lack of norms, standardization, and clinical employment follows. However, instead of eliminating these deficiencies from older APT's, the authors present a new one with seeming ease. Indeed, one APT was done as a hobby and has since been abandoned by its originator.[7] Raskin and Weller (1953) cite two master's theses (Biggs, 1952; Lax, 1953) whose authors developed their own auditory projective tests instead of working with APT's already established.

---

[7] D. R. Stone, personal communication, June 27, 1958.

It is to be hoped that the present paper will bring this situation, as well as awareness of older projective devices, to the attention of psychologists so that earlier tests may be explored before newer ones are thrust upon us. Like the broom of the sorcerer's apprentice, the problems are being multiplied instead of satisfied.

The situation strikes the present writers as resembling that of 15 years ago when many psychologists tried to gain fame by developing a new projective device, or earlier when many physicians sought immortality by endeavoring to describe and name a new phobia. Today the trend seems to be the offering of a new APT suitable for use with the blind.

A comparative investigation should be made of all recorded APT's and the best one, or better parts of several, designated by scientific methods rather than new APT's designed by intuitive cerebration.

Such an investigation should also consider the occurrence of similar sounds in different APT's. For example, train noises and the sound of footsteps are both parts of the stimuli of the Auditory Sound Association Technique, the Sound Apperception Test, and the Auditory Apperception Test (or Wilmer's APT, Bean's APT, and Stone's APT). Seven additional, similar types of sounds occur in both Wilmer's and Stone's APT's. The question of whether such stimuli are duplicates or have different meanings in their respective contexts should be settled. Such factors as the influence of background noises on foreground sounds, and the effect of sequence in the presentation of several sounds should all be investigated before additional APT's are born and then abandoned like unwanted children.

The present writers would suggest that the significance of single sounds be examined first. After that has been determined combinations and sequences could be systematically evaluated. Sound types repeated in two or more APT's might first be investigated. In addition to footsteps, train sounds, and water noises (found in at least three APT's), the list would include groaning, crying, laughing, quarreling, singing, crowd sounds, bird calls, and machine noises. In this manner a sound experimental foundation for APT's would be secured.

A study by Stone might be of especial importance here. For he compared the responses of 220 college students who listened to a record of a bird chirping, a train coming from a distance which comes closer until a crash is heard, and a trumpet fanfare, to the protocols of 74 students who were asked to imagine the sounds which were verbally described to them. It was found that "responses to imagined sounds were either similar to the recorded version, or included them plus additional elaboration. The use of imaginary (verbally described) sounds seems as productive, though less structured than the recorded stimuli" (Stone, 1955, p. 254).

These results indicate that sentences may be substituted for complex sound series and suggest that single sounds may be replaced by single words. If such an hypothesis is tenable then a word association test might indicate responses to imagined unitary sounds. Sounds in brief combination might have their significance determined in the manner of a sentence completion test and noises in more involved arrangements or in descriptive situations might have their interrelationships and sequence determined by Insight Test or VTAT techniques. Since both visual and vocal presentation of the latter stimuli were found to be alike (Lebo, 1959; Sargent, 1956) massive preliminary data might be collected in group situations by means of mimeographed stimuli and written responses.

The apparent ingenuity of such a proposal will not serve for all APT stimuli. How verbally describe some of these unstructured noises of Bean's APT? Nor could unstructured stimuli from 3-DAT projective tests be verbalized for that matter.

The 3-DAT type of test also shows a tendency towards duplication with little use of older and apparently equally good devices. A comparison of these examinations to determine the best one or the best features of several would certainly be in order. Preliminary investigations of such independent features as texture, size, shape, weight, temperature and their influences on responses alone and in combination should be undertaken.

From such a beginning an APT or 3-DAT with the nuances of a Rorschach Psychodiagnostic might eventuate. Scoring methods for the new test may first have to free themselves from an overdependence on TAT methodology. Shakow and Rosenzweig (1940) have developed an original scoring system for the Tautophone that is closely related to that test. They were influenced, however, by the Rorschach ink blot scoring system. One well based APT and 3-DAT should attract a variety of research efforts and empirical accretions which may serve to produce original and subtle scoring procedures over time, unless a giant such as Rorschach appears.

Before embarking upon the formulation of one perfect APT or 3-DAT the necessity of different types of projective devices in various sensory modalities should be determined. If a VTAT in vocal or Braille form is as effective with blind subjects as a TAT with sighted, then elaborate sound tracks and cumbersome three-dimensional figures may not be needed. Nor may it be advisable to approach various sensory doors with unusual devices that ultimately rely on TAT scoring. These tests should all be compared to VTAT findings. If the same kind of information is elicited, less convenient tests might be eliminated. If qualitative differences in APT's and 3-DAT's are smothered by quantitative TAT type approaches then this familiar, but perhaps inappropriate, scoring procedure should be eschewed for such tests.

Investigations comparing the effectiveness of several personality measurements appealing to different sense organs are few in number and usually impressionistic in nature. Thus, Kerman[8] stated that percepts gained by tactual exploration of cypress knees by the blind were practically identical with those found through visual inspection. The superiority or inferiority of the KCK in relation to other methods has not been established. Twitchell-Allen (1947) found the 3-DAT to be more informative in certain aspects of personality than the Rorschach test when results were compared in the same individual. Information obtained from the APT was found to be "at least comparable to that obtained from the TAT" (Braverman & Chevigny, 1955). The scant results are, with the exception of Twitchell-Allen's, discouraging. A criterion measure should be established (the present writers would like to propose a VTAT for this position) and all other tests measured against it. If they do not secure more information, or a different kind of information, they may be shelved for special usage.

This special usefulness depends upon their being rarely encountered. Because of this aspect they may serve as projective reevaluators of test wise subjects or institutional habitues. Even such patients will probably not be familiar with many of the tests herein described and may find themselves responding, once again, to unstructured stimuli.

G. SUMMARY

Projective stimuli recommended or used as suitable evaluative procedures with blind persons have been presented. Because many of them are little known they were discussed in detail. While the number of such methods is still limited (approximately 20 were found in the literature) noticeable development has occurred. They seem to be increasing in a dangerous manner. That is, instead of older tests being investigated newer ones are being devised.

Because of similarities of names in many such devices a standardized titling procedure was suggested. Because of similarities of content a fundamental research program was suggested. It was indicated that much of the experimentation on the dynamic meaning of sound stimuli could justifiably be conducted by associations to written words or sentences. It was also pointed out that possible nuances of approach may be eliminated by dependence upon a TAT scoring system. Devices that were developed because they were different would seem to merit original scoring procedures.

---

[8] E. F. Kerman, personal communication, May 26, 1958.

It was suggested that one immediate use for lesser known auditory or three-dimensional projective tests would be the diagnosis of test wise, sighted clients.

REFERENCES

1. ALLEN, R. M. Personality Assessment Procedures. New York: Harper, 1958.
2. ANASTASI, A. Psychological Testing. New York: Macmillan, 1957.
3. ASHFORD, J. T., & THURMAN, A. J. The role music can play in child life. *New Outlook for the Blind,* 1955, **49**, 98-101.
4. BAUMAN, M. K., & HAYES, S. P. A Manual for the Psychological Examination of the Adult Blind. New York: Psychological Corp., 1950.
5. BEAN, K. L. Development of a sound apperception test. *Amer. Psychol.*, 1957, **12**, 390. (Abstract)
6. BENDER, L., & WOLTMANN, A. G. The use of plastic material as a psychiatric approach to emotional problems in children. *Amer. J. Orthopsychiat.*, 1937, **7**, 283-300.
7. BIGGS, J. I. An experiment in the use of a projective technique with blind children. Unpublished master's thesis, Univ. Texas, 1952.
8. BRAVERMAN, S., & CHEVIGNY, H. The Auditory Projective Test. New York: Amer. Found. Blind, 1955. (Mimeographed)
9. BRIELAND, D. Personality problems of the blind and visually handicapped as revealed by a projective technique. *Amer. Psychol.*, 1950, **5**, 340. (Abstract)
10. BRIGGS, D. L., GAEDE, D. C., & WILMER, H. A. Projective sound testing of emotionally ill patients. *U. S. Armed Forces Med. J.*, 1956, **7**, 1764-1770.
11. DAVIDS, A. Some personality and intellectual correlates of intolerance of ambiguity. *J. Abn. & Soc. Psychol.*, 1956, **51**, 415-420.
12. DAVIDS, A., & MURRAY, H. A. Preliminary appraisal of an auditory projective technique for studying personality and cognition. *Amer. J. Orthopsychiat.*, 1955, **25**, 543-554.
13. DEAN, S. I. Adjustment testing and personality factors of the blind. *J. Consult. Psychol.*, 1957, **21**, 171-177.
14. FAIRBAIRN, W. R. D. An Object-Relations Theory of the Personality. New York: Basic Books, 1954.
15. FASSETT, K. K. A preliminary investigation of the Sargent test. *J. Clin. Psychol.*, 1948, **4**, 45-55.
16. GOODENOUGH, F. L., & HARRIS, D. B. Studies in the psychology of children's drawings: II, 1928-1949. *Psychol. Bull.*, 1950, **5**, 369-433.
17. GRINGS, W. W. The verbal summator technique and abnormal mental states. *J. Abn. & Soc. Psychol.*, 1942, **37**, 529-545.
18. HARRIS, W. W. Notes on initial experiments with bas relief projective material for blind subjects. *Rorschach Res. Exch.*, 1947, **11**, 80-81.
19. ———. A bas relief projective technique. *J. of Psychol.*, 1948, **26**, 3-17.
20. JUNG, C. G. Studies in Word Association. London: Heinemann, 1918.
21. KADIS, A. L. Finger-painting as a projective technique. *In* L. E. Abt & L. Bellak (*Eds.*). *Projective Psychology.* New York: Knopf, 1952. (Pp. 403-431.)
22. KERMAN, E. F. Cypress knees: A new three-dimensional projective technic (KCK). *Psychiat. Quart. Suppl.*, 1957, **31**, 1-10.

23. ———. Cypress knees and object-relations. *Psychiat. Quart.*, in press. (a)

24. ———. Psychodiagnosis and psychodynamics from an object-relation frame of reference. *Psychiat. Quart.*, in press. (b)

25. Korhonen, G. V. Music as an educational value for the blind. *New Outlook for the Blind,* 1956, **50**, 19-94.

26. Kunze, K. R. Musical Reverie Test. *In* H. A. Murray (*Ed.*), *Explorations in Personality*. New York: Oxford Univ. Press, 1938. (Pp. 550-552.)

27. Lax, B. A. A comparison of the imaginative productions of congenitally blind and seeing children to structured auditory stimulation. Unpublished master's thesis, Univ. Texas, 1953.

28. Lebo, D. Immediate affective reaction to TAT cards. *J. Clin. Psychol.,* 1955, **11**, 297-299.

29. ———. Scoring the verbal TAT, with illustrative protocols. *In* (P. J. Sherry), *A Comparison of Responses to Written and Spoken TAT Card Descriptions.* Unpublished master's thesis, Richmond Prof. Inst., 1959. (a)

30. ———. An empirical approach to problems concerning the diagnostic value of a pictureless TAT. *J. Proj. Tech.,* 1959, **23**, 107. (b)

31. Lebo, D., & Harrigan, M. Visual and verbal presentation of TAT stimuli. *J. Consult. Psychol.,* 1957, **21**, 339-342.

32. Lebo, D., Noblin, C. D., & Toal, R. A. Reliability and validity of the Manifest Anxiety Scale with schizophrenic patients and prison inmates. *Amer. Psychol.,* 1957, **12**, 416. (Abstract)

33. Lebo, D., & Sherry, P. J. Visual and verbal presentation of TAT descriptions. *J. Proj. Tech.,* 1959, **23**, 59-63.

34. Lehmann, H. E., & Risquez, F. A. The use of finger paintings in the clinical evaluation of psychotic conditions: A quantitative and qualitative approach. *J. Ment. Sci.,* 1953, **99**, 763-777.

35. Levine, J. Research on a three-dimensional apperception test; a new projective technique. Personal communication, cited by W. W. Morris. A survey of projective techniques for use with the blind adult. *In* W. Donahue & D. Dabelstein (*Eds.*), *Psychological Diagnosis and Counseling of the Adult Blind.* New York: Amer. Found. Blind, 1950.

36. Lowenfeld, V. The Nature of Creative Activity. New York: Harcourt, Brace, 1939.

37. ———. Creative and Mental Growth. New York: Macmillan, 1955.

38. McAndrew, H. The use of projective techniques in the personality evaluation of the blind. *Amer. Psychol.,* 1950, **5**, 340. (Abstract)

39. Melville, H. Mardi and a Voyage Thither. London: Constable, 1922.

40. Morris, W. W. A survey of projective techniques for use with the blind adult. *In* W. Donahue & D. Dabelstein (*Eds.*), *Psychological Diagnosis and Counseling of the Adult Blind.* New York: Amer. Found. Blind, 1950. (Pp. 114-129.)

41. Murray, H. A. Thematic Apperception Test Manual. Massachusetts: Harvard Univ. Press, 1943.

42. Napoli, P. J., & Harris, W. W. Finger painting for the blind. *J. of Psychol.,* 1948, **25**, 185-196.

43. Raskin, N. J., & Weller, M. F. Current Research in Work for the Blind. New York: Amer. Found. Blind, 1953.

44. Richter, J. P., *Ed.* The Literary Works of Leonardo da Vince. (2nd ed.) New York: Oxford Univ. Press, 1939. (2 vols.)

45. Rosanoff, A. J. Manual of Psychiatry and Mental Hygiene. (7th ed.) New York: Wiley, 1938.

46. Rotter, J. B. Incomplete Sentences Blank—Adult Form. New York: Psychological Corp., 1950.

47. ROTTER, J. B., & RAFFERTY, J. E. Manual: The Rotter Incomplete Sentences Blank. New York: Psychological Corp., 1950.

48. ROTTER, J. B., & WILLERMAN, B. The incomplete sentences test as a method of studying personality. *J. Consult. Psychol.*, 1947, **11**, 43-48.

49. SARGENT, H. D. An experimental application of projective principles to a paper and pencil personality test. *Psychol. Monog.*, 1944, No. 5.

50. ———. The Insight Test: A Verbal Projective Test for Personality Study. New York: Grune & Stratton, 1953.

51. ———. Insight Test prognosis in successful and unsuccessful rehabilitation of the blind. *J. Proj. Tech.*, 1956, **20**, 429-441.

52. SEASHORE, C. E., & LING, T. L. Comparative sensitiveness of the blind and seeing persons. *Psychol. Monog.*, 1918, No. 2.

53. SHAKOW, D., & ROSENZWEIG, S. The use of the Tautophone ("Verbal Summator") as an auditory apperception test for the study of personality. *J. Personal.*, 1940, **8**, 216-226.

54. SKINNER, B. F. The Verbal Summator and a method for the study of latent speech. *J. of Psychol.*, 1936, **2**, 71-107.

55. STEIN, M. I. The use of a Sentence Completing Test for the diagnosis of personality. *J. Clin. Psychol.*, 1947, **3**, 46-56.

56. STONE, D. R. A recorded Auditory Apperception Test as a new projective technique. *J. of Psychol.*, 1950, **29**, 349-353.

57. ———. The Auditory Apperception Test. Beverly Hills: Western Psychological Services, 1953.

58. ———. Responses to imagined auditory stimuli as compared to recorded sounds. *J. Consult. Psychol.*, 1955, **19**, 254.

59. SYMONDS, P. M. The Sentence Completion Test as a projective technique. *J. Abn. & Soc. Psychol.*, 1947, **42**, 320-329.

60. TRESSELT, M. E., & LEEDS, D. S. The frequencies of responses by males and females (N = 124) for each of the 100 Kent-Rosanoff stimulus words, ages 22-25. *Psychol. Newsletter*, 1953, **5**, 39-74. (a)

61. ———. The responses and frequencies of responses for males and females (18-21) to the Kent-Rosanoff list. *Psychol. Newsletter*, 1953, **5**, 1-36. (b)

62. TRUSSELL, M. A. The diagnostic value of the verbal summator. *J. Abn. & Soc. Psychol.*, 1939, **34**, 533-538.

63. TWITCHELL-ALLEN, D. A Three-Dimensional Apperception Test: a new projective technique. *Amer. Psychol.*, 1947, **2**, 271-272. (Abstract)

64. ———. A Three-Dimensional Apperception Test. New York: Psychological Corp., 1948.

65. ———. The significance of movement in diagnosis and therapy; A presentation of the Three-Dimensional Apperception Test. *In* F. Baumgarten (*Ed.*), *La psychotechnique dans le monde moderne*. Paris: Presses Univ. de France, 1952. (Pp. 212-214.)

66. WILMER, H. A. An auditory sound association technique. *Science*, 1951, **114**, 621-622.

67. WILMER, H. A., & HUSNI, M. A. An auditory projection technique for psychiatric research. *J. Lancet*, 1952, **72**, 17-24.

68. ———. The use of sounds in a projective test. *J. Consult. Psychol.*, 1953, **17**, 377-383.

69. WOLFF, W. Diagrams of the Unconscious. New York: Grune & Stratton, 1948.

Child Guidance & Speech Correction Clinic
625 Ocean Street
Jacksonville 2, Florida

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 39-41.

STUDIES IN CONFORMITY BEHAVIOR: A METHODOLOGICAL NOTE*

Department of Psychology, University of Michigan

LEO LEVY[1]

A. PROBLEM

The extent to which an individual will allow group pressure to affect his reported solution of a simple perceptual task has been the subject of a continuing series of experiments by Asch (1951, 1952, 1956). Crutchfield (1955), using the same basic situation devised by Asch, constructed a laboratory in which subjects are placed in individual booths and presumably receive information from other subjects via an electrical panel. The question arises as to the comparability of the Asch and Crutchfield situations. Are subjects who give their answers verbally, face-to-face with a group, exposed to the same degree of pressure to conform as those placed in a semi-anonymous booth? A study which bears directly on this problem has been carried out by Deutsch and Gerard (1955). They found a very significant reduction in the number of conformity responses from the Asch to the Crutchfield situation.

B. THE PRESENT STUDY: METHOD

For purposes of the present study apparatus similar to that used by Crutchfield was constructed. The problems used were identical with Asch's, each problem offering a standard and three comparison lines drawn to his specifications. The problems were projected on a screen in front of the subjects. A total of 19 trials were used. The first trial was a neutral demonstration trial followed by 18 test trials (two blocks of nine each) as used by Asch. One hundred and two undergraduate male subjects, drawn from lower division psychology courses, were tested five at a time. They were seated in adjacent booths and acquainted with the type of problem with which they would be dealing. They were led to believe that each person would occupy a different position in the group as indicated by a light on the left of their panels. They were instructed to answer in order, waiting for answers to flash on their

---

* Received in the Editorial Office on March 7, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This report is part of a doctoral dissertation submitted to the department of psychology at the University of Washington, Seattle, Washington. The thesis was sponsored by Dr. A. L. Edwards.

panels from the men who occupied earlier positions in the group before giving their own. The panels, aside from the column of lights indicating the subjects' turn, consist of three columns of lights with a switch underneath each. These lights correspond to the three comparison lines in each problem and the answer was indicated by depressing the appropriate switch. In so doing, the subject was informed, he lit the appropriate light on his panel and on all the others, thus communicating his answer to the group. Switches used for the panels were noiseless so that one subject would not know when another was pressing a switch. The apparatus was so constructed that each subject controlled only one row of lights on his own panel corresponding to his position in the group. The remainder of the lights were controlled by a master panel operated by the experimenter.

The subjects were first tried in Position 3, then in Position 2, and finally in Position 5 which they retained for the remainder of the trials. Subjects' choices were shown on the experimenter's panel and recorded. Of the 19 trials given, 12 were critical trials where the subject found himself preceded by an unanimous incorrect majority. At the conclusion of the nineteenth trial, the subjects were asked to write down what they thought had happened. Following this the experiment was explained to them, and they were requested not to discuss the experiment with their classmates.

C. RESULTS

Of the 102 subjects tested in the conformity situation, 96 usable records were obtained. The subjects ranged in age from 17-38 years of age with a median age of 20 and were of heterogeneous racial extraction.

Conformity responses accounted for 19.10 per cent of the total number of responses. This is a striking reduction from Asch's reported 36.80 per cent and is in better agreement with the figure of 23.10 per cent computed from the data of Deutsch and Gerard from a small sample using a similar experimental situation.

Written reports taken after the test runs were checked to see how many subjects suspected that the situation was rigged. Subjects who listed this hypothesis either alone or among several other possible explanations for the discrepancy between their perceptions and the information fed to them are referred to as "skeptics." Of the total sample 35.42 per cent fall in this group. This figure refers to those persons who were sufficiently convinced of their hypothesis to write down their suspicion at the conclusion of the experiment. It may be safely assumed that others entertained this notion, but not with sufficient conviction to express it. It may also be assumed that within the one-third of the sample who expressed their suspicion, all degrees of certainty are present, ranging from the subject who definitely and tersely states that the situation was controlled by the experimenter to the subject who tentatively suggests this, perhaps enmeshed among other popular alternative explanations.

One additional contrast of the present data with Asch's results pertains to the stability of the conformity response. Asch, using two identical blocks of nine trials, found no significant decrease in conformity responses from Block 1 to Block 2. In the present study, a significant decrease in conformity responses from Block 1 to Block 2 was observed ($p < .05$). A further examination of the data indicates that this is a general effect; i.e., skeptics show no greater tendency to decline in the number of conformity responses given during the second block of trials than non-skeptics. Of the 59 subjects who demonstrated conformity behavior at all (i.e., conformity score greater than zero), 32 showed a decrease, 15 remained stable, and 12 showed an increase in conformity responses from Block 1 to Block 2.

D. Discussion and Summary

The present study indicates that the laboratory situation created by Crutchfield is far less effective than the original face-to-face situation created by Asch. Three sources of evidence are offered in support of this conclusion. The first concerns the drop in conformity responses; a reduction previously noted by Deutsch and Gerard. Second, a large number of subjects expressed some suspicion that the situation was rigged. The number of skeptics found represents a considerable increase over the figures reported by either Asch or Crutchfield. Finally, the tendency to conform is not found to be a stable effect in the present study, decreasing significantly in the second block of trials.

References

1. Asch, S. E. Effects of group pressure upon the modification and distortion of judgment. *In* Guetzkow, H. (*Ed.*) *Groups, Leadership, and Men.* Pittsburgh: Carnegie, 1951.
2. ———. Social Psychology. New York: Prentice-Hall, 1952.
3. ———. Studies of independence and conformity: 1. A minority of one against a unanimous majority. *Psychol. Monog.*, 1956, **70**, No. 9 (Whole No. 416).
4. Crutchfield, R. Conformity and character. *Amer. Psychol.*, 1955, **10**, 191-198.
5. Deutsch, M., & Gerard, H. A study of normative and informational social influences upon individual judgment. *J. Abn. & Soc. Psychol.*, 1955, **51**, 629-636.

Department of Psychology
University of Michigan
Ann Arbor, Michigan

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 43-52.

VARIABLY BLURRED PROMPTING: I. METHODOLOGY AND APPLICATION TO THE ANALYSIS OF PAIRED-ASSOCIATE LEARNING*

Department of Psychology, Harvard University

M. L. Israel[1]

A. Introduction and Problem

A single verbal response may have several sources of strength (5, ch. 9-11). A case which lends itself to investigation is this: one or more sources are present, but are not sufficient to evoke the response; a supplementary source is made available to *S*, who varies it along a continuum of increasing effectiveness until the response is evoked. It has been suggested (5, p. 228) that such sources combine additively: the greater the contribution to probability of emission made by sources already present ("principal" sources), the smaller the contribution a supplementary source needs to add. A measurement of the contribution from a supplementary source may thus be used as an indirect measure of the contribution from principal sources.

This paper reports a German-English paired-associate experiment in which whenever *S* was unable to recall the English member of a pair, he clarified a blurred printed image of that member until the image prompted a correct response. The blurred image is a supplementary source of strength because of its similarity in form to a clearly printed version of the word; for this reason it may be termed a "textual" (5, p. 65) source. A measure of the amount of contribution required from this source to prompt a correct response is operationally equivalent to a "recognition threshold" and will be designated as such.

The experiment contains at least two principal sources of strength for correct responses. Each time *S* goes through the full set of associates, all the response (and stimulus) members are made more probable in his repertoire. Each encounter with the full set "defines" for *S* the set of relevant words. This may make the experimental situation itself a principal source.

---

* Received in the Editorial Office on March 11, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was carried out during the tenure of a Predoctoral Fellowship from the National Institute of Mental Health, United States Public Health Service, and was also supported by grants from the Carnegie Corporation, the Ford Foundation, and the National Science Foundation. The author is especially indebted to Wells Hively, among several colleagues who read the manuscript.

The heightened probability of response members due to this source may be termed "response-member" learning to distinguish it from the "intraverbal" (5, p. 71) learning which takes place at the same time and which creates a second principal source. In intraverbal learning, a German word acquires discriminative control over the emission of the correct English equivalent. Each German word becomes a source of strength for its correct equivalent, and functions as such each time it is presented to *S*.

The presence of principal sources will cause recognition thresholds to decrease by causing an increase in response probability. Thresholds may also be affected by an improvement, over the course of the experiment, in *S*'s ability to discriminate all blurred stimuli. This would contribute to threshold drop not by adding to response probability, but by shifting the stimulus control of response probability toward increasingly blurred values of any textual stimulus.

The present experiment assesses the contribution made by the intraverbal source alone. The combined effects of response-member learning and blur-discrimination learning are evaluated by interleaving "control" trials into the basic set of German-English ("experimental") trials.

The present study bears on the controversy over whether paired-associate learning is all-or-none in nature (2, 3, 4). An *S* is presented with a paired-associate pair and given an opportunity to study it. At some later point we measure the contribution the first member makes to the probability of emission of the second member. Is the contribution always 100 per cent (a correct recall) or 0 per cent? Or can it assume some intermediate value—i.e., do partial associations exist?

Consider a pair whose response member is not recalled until the pair's fourth presentation. (*a*) Does the amount of prompt required on the second and/or third presentation indicate that the stimulus member was making some partial contribution to response probability at either or both points? This question is identical with the one Rock raises as to whether "items . . . which are not [yet] learned . . . leave behind any increment of associative strength" (3, p. 70). (*b*) If there is a partial contribution, is it greater on the third presentation than on the second? Traditional association theory suggests that associations grow in strength up to the point of recall.

B. METHOD[2]

1. *Experimental Trials*

A German word was projected onto a fixed translucent screen, about 51 cm. from S's eyes, and remained there throughout the trial, always in perfect focus. The S attempted to give the correct English equivalent, and was given as much time as he wished. If he could not do so (necessarily the case on the first presentation of each German word) he depressed a telegraph key, which simultaneously (*a*) caused a shutter to expose the blurred image of the English equivalent on a movable screen, and (*b*) closed a circuit (as long as the key was depressed) to a constant-speed motor which drove the movable screen from a plane at which the intercepted image of the English equivalent was extremely blurred toward a plane at which the image was in perfect focus.

---

[2] A more detailed account of the apparatus and a report of two subsequent related experiments is in preparation as part of a doctoral dissertation.

The S released the key when the image became clear enough to prompt a response. Release of the key simultaneously (*a*) stopped clarification, and (*b*) re-covered the blurred image with its shutter. The S then announced his answer. If correct, E exposed a clear image of the English word in a position just above the German word on the fixed screen. The S was then given as much time as he wished to study the German-English pair. (This interval will be termed the "study period.") If S was incorrect, he was informed of this and S clarified the image some more and made another attempt at the correct response. Total clarification required for a correct response constituted the threshold measurement. Clarification was quantified in terms of the screen's displacement in cm. from its constant initial position. Screen speed was 6.21 mm/sec.

After studying the pair, S operated a second control switch that (*a*) reset the movable screen and (*b*) operated a slide changer that projected the images for the next trial. (All images for a trial were projected from a single slide which transilluminated both the fixed and the movable screen.) Whenever S was able to make a correct response without using the prompter ("unprompted recall"), E exposed the correct equivalent in clear focus on the fixed screen; S studied the pair if he wished to, and advanced to the next trial. (Unprompted recalls were only possible on second and later presentations of the German words.)

2. *Control Trials*

The procedure duplicated that of experimental trials, but with the following differences. No German word appeared on the fixed screen. Subject's task was to clarify a blurred image of an English word until he could just barely recognize it. After a correct recognition, a clear image of the English word appeared on the fixed screen, but no study period ensued; S simply advanced to the next trial.

3. *Verbal Materials*

A large number of 8-letter English words were selected at random from a dictionary. Unusual words (in *E*'s judgment) were thrown out. From the remaining pool of words, 20 were drawn randomly for use on control trials. Another 20 were drawn for which suitable German equivalents could be found in an English-German dictionary. "Suitability" required: (*a*) that the German equivalent be within 5 and 12 letters long (median length was 7); (*b*) that it also be neither a cognate nor a word whose meaning was likely to be known to the *S*s; and (*c*) that no obvious mnemonic device for associating it with its English equivalent be readily available, in *E*'s judgment.

4. *Arrangement and Presentation of Slides*

Twenty experimental and 20 control slides were presented in alternation. Presentation of the full set of 40 slides will be termed a "cycle." A different random order of slides was used on each cycle (with the restriction that there was always an alternation of experimental and control slides). The number of cycles presented depended on *S*'s rate of learning: if *E* judged, on the basis of *S*'s performance, that five or fewer thresholds would be measured on the next cycle (i.e., 15 or more unprompted recalls would occur) the experiment was terminated.

Median cycle time was 25 minutes, range 18-38. Time between cycles was about five minutes, during which *S* was supplied with reading matter. Median number of cycles per *S* was 3, range 2-5. Median length of experiment (including the instruction period) was 90 minutes, range 70-140.

5. *Subjects*

Eighteen students (median age 20) at local schools and universities were recruited through a university student employment office. Remuneration was hourly. Subjects were restricted to those claiming (*a*) 20/20 acuity (corrected or not) and (*b*) little or no knowledge of German. Seven *S*s were subsequently tested for binocular near vision on a Freeman Acuity Tester and all proved to be 20/20 or better. Experimental design was such that previous knowledge of German could not affect the variables under principal investigation.

C. Results

1. *Existence of Partial Associations*

Data for subject L.I. are chosen for illustrative analysis since her mean thresholds were similar to the means for pooled data of all *S*s. Thresholds for each experimental and control word are shown in *A* and *B* of Figure 1. Mean thresholds are shown in Figure 2A.

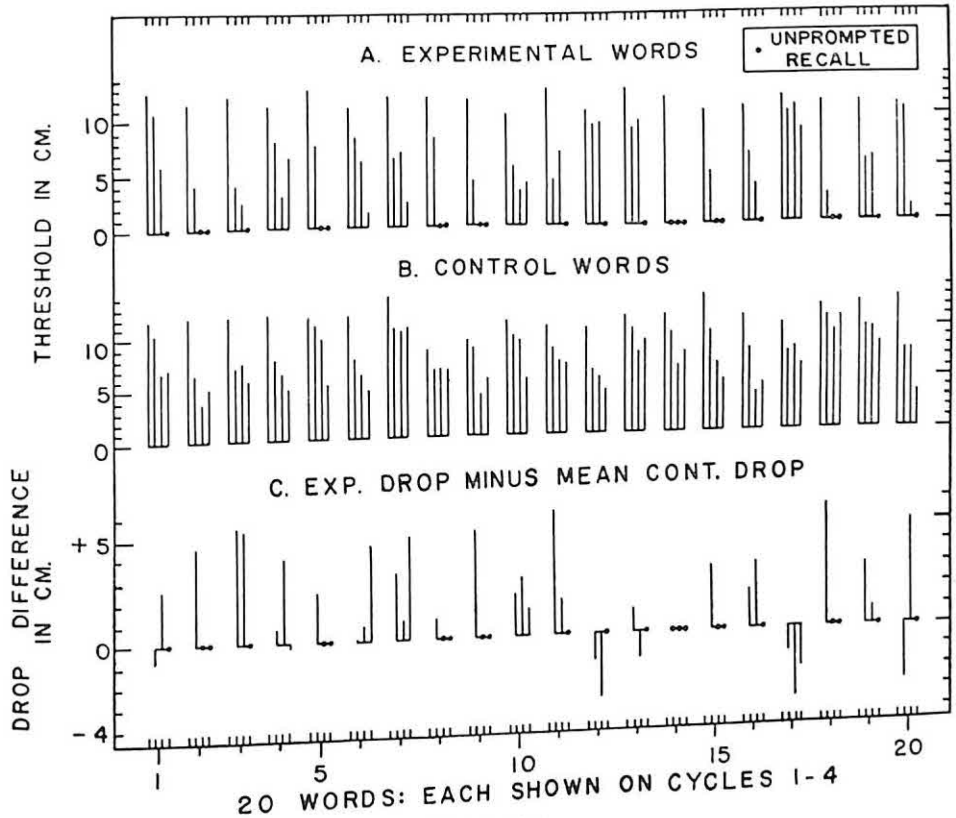

FIGURE 1

DATA FOR SUBJECT L.I.

A and B: Four lines for each word represent, from left to right, thresholds on Cycles 1-4.

C: Each experimental word is taken separately, and the drops of its second-, third-, and fourth-cycle thresholds are calculated. From each of these drops is subtracted the mean drop of all 20 control thresholds for the cycle in question. The three resulting drop differences are lined up vertically with the second-, third-, and fourth-cycle thresholds of the experimental word in Figure 1A on which the differences are based.

A decrease in a threshold from its first-cycle value will be termed a "threshold drop" or simply a "drop." Drop in control thresholds is attributable to response-member and blur-discrimination learning. Experimental drop is caused by these factors and also by the effect of German words as intraverbal sources of strength. The best estimate of the intraverbal contribution alone is the difference between experimental and control drops—the "drop difference." Subject L.I.'s drops are plotted in Figure 2B; the drop differences are shown by the distances between experimental and control points.

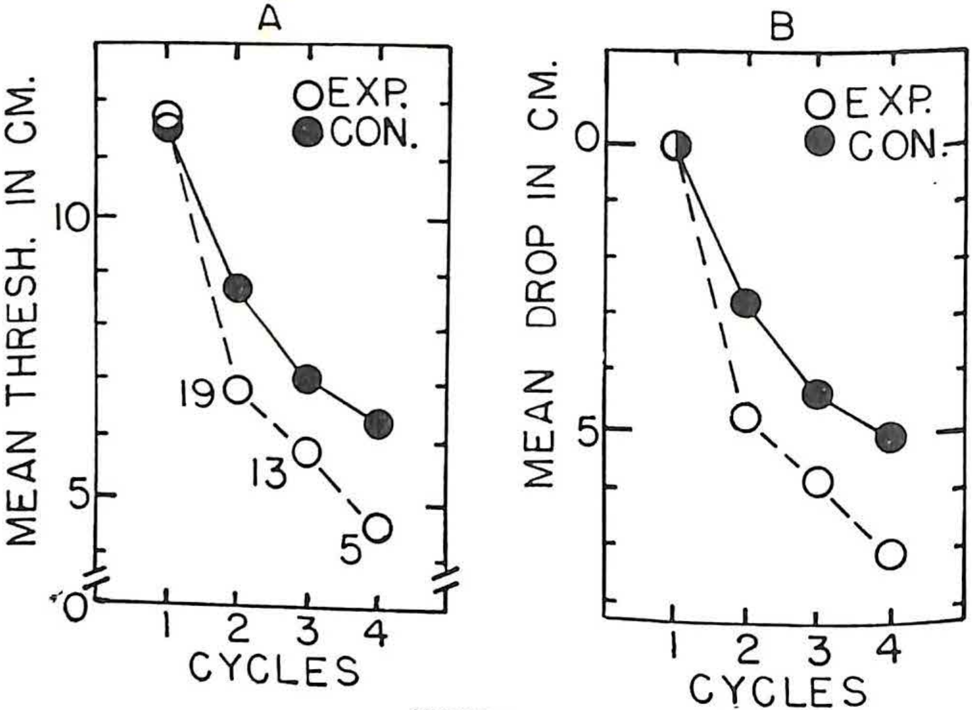

FIGURE 2

MEAN CURVES FOR SUBJECT L.I.

A: Mean thresholds. Control means on all cycles, and experimental means on Cycle 1 are based on 20 thresholds. On second and later cycles, unprompted recalls are *not* included as "prompts" of 0.0 cm. when calculating means. The number of thresholds entering experimental means on second and later cycles is indicated by numerals adjacent to the plotted experimental points.

B: Drop of experimental and control mean thresholds from their first-cycle values.

Figure 3 shows mean threshold and mean drop curves for all 18 *S*s arranged alphabetically by first initial. Only two *S*s—C.O. and S.H.—failed to show consistently larger experimental drops. W.A. is questionable evidence; since her Cycle 1 experimental mean was considerably larger than her Cycle 1 control mean, it may be argued that her experimental values had that much more distance over which they could and did drop. Curves for pooled data of all subjects are also shown. These suggest that roughly 70 per cent of the experimental drop is attributable to the same factors which cause the control drop.

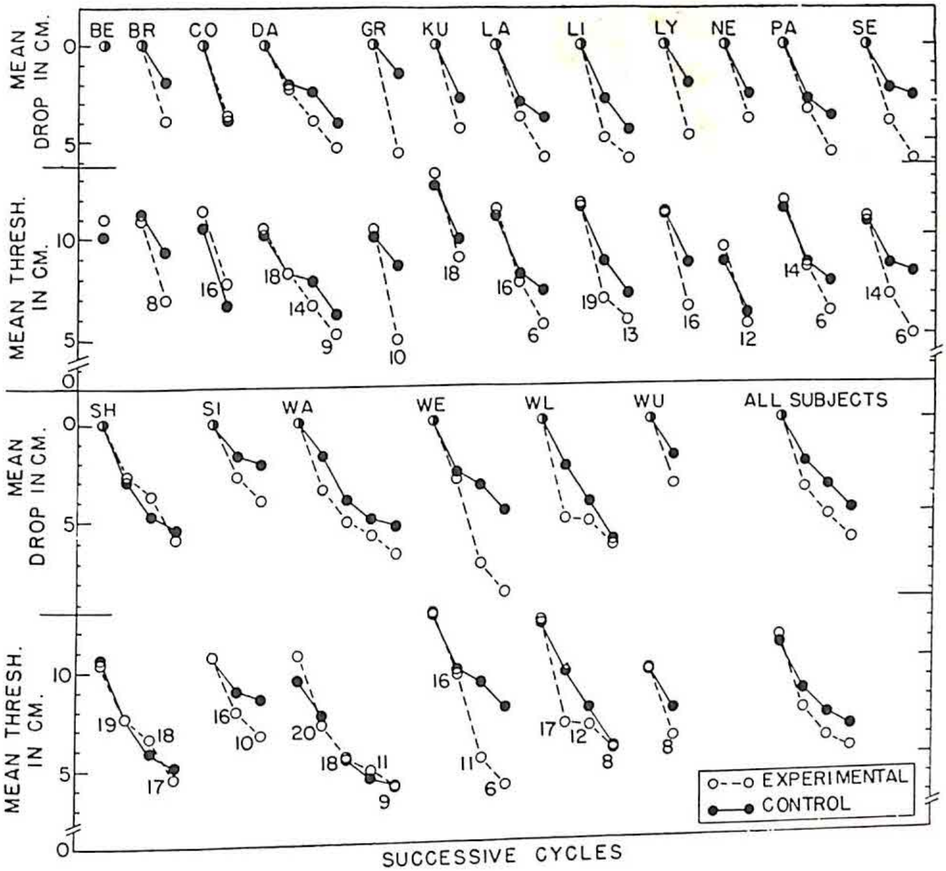

FIGURE 3

Mean Curves for Each Subject and for Pooled Data

Mean threshold curves are directly below mean drop curves for the same *S*. In curves for individual *S*s, a cycle is not plotted if the experimental mean for that cycle is based on five or fewer thresholds. In curves for pooled data (lower right) the fifth cycle is not plotted, being based on only 14 thresholds. See Figure 2A for explanation of numerals adjacent to plotted points.

2. *Changes in Strength Prior to Unprompted Recall*

In Figure 1C, the drop difference measure is applied to individual experimental words of L.I. to obtain an estimate of the contribution of each German word to probability of response on each cycle. There is no clear trend in the drop differences. The variability displayed is typical of other subjects.

The eight words recalled by L.I. for the first time on Cycle 4 were pooled. The drop differences on Cycle 3 were not significantly larger than those on Cycle 2 [Wilcoxon signed ranks (1)]. All words recalled for the first time on Cycle 4 were then pooled over all *S*s (except C.O. and S.H., whose data were omitted because their experimental drops failed to be consistently larger than their control). For the 30 words in this pool, the drop differences on Cycle 3 were still not significantly larger than those on Cycle 2, using the same test.

All words whose thresholds could be followed through three cycles (88 words in all)—regardless of whether they were recalled on Cycles 4 or 5, or were unrecalled throughout the experiment—were then pooled over all *S*s (omitting C.O. and S.H.). A one-tailed sign test (1) showed the drop differences on Cycle 3 (mean 2.0 cm.) to be significantly larger than those of Cycle 2 (mean 0.9 cm.) at the .005 level of confidence. (In a subsequent experiment,[3] however, a similar comparison failed to yield a significant difference between drop differences of Cycles 2 and 3.)

Pooling over the same *S*s, only 24 words could be followed through four cycles. A Friedman two-way analysis of variance by ranks (1) showed no significant differences among the drop differences of Cycles 2, 3, and 4.

D. DISCUSSION

An intermediate probability of response can manifest itself (*a*) by affecting the amount of additive prompt necessary for a prompted recall or (*b*) in an actual emission of the response ("unprompted recall"). Subjects were allowed an unlimited time (the "pre-prompt" period) for attempting unprompted recalls in the hope that most if not all of such recalls would emerge at that time and that few if any would emerge during the threshold measurement itself. (This explains why it was impractical to regulate the length of the study period allowed at the end of experimental trials. If the study period had been limited, *S* would simply have used the unlimited pre-prompt period to rehearse pairs presented on preceding trials.)

To *E*'s knowledge, no unprompted recalls occurred during the threshold measurements of the present experiment. One did occur in a subsequent experiment and one occurred in a demonstration for a colleague. Both occurred at unusually early points in the clarification process and were striking enough to the *S*s that they remarked upon them to *E*. The possibility of the intrusion of such recalls into the recognition measure itself is a weakness of the present technique and suggests that an additional type of control trial should be run to determine whether such recalls do infect the threshold measure.

A second possible defect is that on control trials there was nothing comparable to the study period activities that *S* engaged in on experimental trials. Rehearsal during the study periods may generate an increase in response-member probability that is separate from, although occurring at the same time as, intraverbal learning. If so, such a separate increase in probability may account for part of the obtained drop differences. A report[4] of an experiment that tested this (and supported the present findings) is in preparation.

---

[3] See Footnote 2.

With the above reservations, present findings are these. First, German words as associative stimuli in a paired-associate learning task do contribute to the probability of emission of their paired English associates on trials on which *S*s are as yet unable to make correct unaided recalls. Partial associations, as defined above, do exist.

Second, the results do not show clear evidence of a cycle-to-cycle increase in this intraverbal contribution prior to the point of recall. The present technique may not be sensitive enough to detect changes of this nature; or, there may in fact be a real fluctuation in strength prior to recall. The expectation of a cycle-to-cycle increase is based on the fact that each cycle gives *S* an additional opportunity to study each pair. But an opportunity for study does not guarantee that such study will in fact occur. Furthermore, an *S* can and does rehearse a pair at times other than the study period of the trial on which that pair was presented.

Blurring is but one of a myriad of possible prompting techniques. Every technique for measuring a recognition threshold is a candidate to be a variable-prompting measure. Investigations with other techniques will be necessary to crosscheck the results reported here, and could turn up a technique of greater sensitivity.

If further investigation confirms the existence of partial associations, several interesting applications of variable prompting suggest themselves. In the field of verbal learning, we may be able to define new measures of learning. Traditional "error" or "per cent correct" curves measure complete associations only. The "drop differences" of the present experiment are an attempt to measure partial associations. A measure combining both types is also possible. For example, if the experimental drop curve of Figure 2B had been derived from experimental mean thresholds in which unprompted recalls had been allowed to enter in the guise of "prompts" of 0.0 cm. in size, the distance between the points of Figure 2B would have been a measure of complete as well as partial associations, instead of partial associations only.

In the field of teaching machines, a variable prompter may be useful for at least two reasons: (*a*) Where partial associations exist but are insufficient to enable recall, allowing a student to turn to a full textual or "echoic" (5, p. 55) stimulus for the answer may be a "waste" of whatever partial intraverbal tendencies have been created. A prompter, by putting partial tendencies to work, may increase the rate at which unprompted recalls emerge. (*b*) A prompter may have important motivational properties. Several *S*s in the present experiment confessed embarrassment over taking money for participating. They had found it fun.

---

[4] See Footnote 2.

E. Summary

Subjects were given a paired-associate learning task. Whenever *S* could not recall the English member of a pair, he clarified a blurred image of that word until a correct response was prompted. A measure of clarification required was used to analyze association formation. Control trials evaluated the effect of extraneous learning factors upon the clarification measure. Results demonstrate the existence of partial associations, but do not show clear evidence for a growth in their strength prior to the point at which unaided recall is possible.

References

1. Mosteller, F., & Bush, R. R. Selected quantitative techniques. *In* Lindzey (*Ed.*), *Handbook of Social Psychology:* Vol. 1. *Theory and Method.* Cambridge: Addison-Wesley, 1954. (Pp. 289-334.)
2. Rock, I. The role of repetition in associative learning. *Amer. J. Psychol.*, 1957, **70**, 186-193.
3. ———. Repetition and learning. *Sci. Amer.*, 1958, **199** (2), 68-72.
4. Rock, I., & Heimer, W. Further evidence of one-trial associative learning. *Amer. J. Psychol.*, 1959, **72**, 1-16.
5. Skinner, B. F. Verbal Behavior. New York: Appleton-Century-Crofts, 1957. Pp. 478.

Department of Psychology
Memorial Hall
Harvard University
Cambridge 38, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 53-57.

A FURTHER FACTOR IN DETERMINING NEARNESS AS A FUNCTION OF LATERAL ORIENTATION IN PICTURES*

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND DORIS C. DEHARDT

A. PROBLEM

Using a standard psychophysical procedure, it has been established (1, 2) that items in the left hand portion of a pictorial scene (photographic print) appear closer than the same items in the right hand portion. For one form of target material, the scene and its mirror image were involved. For another form, a scenic print was variously cropped so as to place the focally attended item in either the left or the right hand parts of the print, the amount of lateral eccentricity being varied as an additional factor. Having proceeded this far, the senior author was left with the general conclusion stated above, but eventually this seemed too general. The question of whether this finding would hold true for background (perceptually distant) items as well as for those in the foreground presented itself. Need for obtaining data to answer this question gave rise to the present investigation.

B. APPARATUS AND PROCEDURE

Two photos of realistic everyday scenes were used. They were somewhat asymmetrical in composition, though one scene more than the other [a fact previously established by Adair and Bartley (1), since the two scenes used in this study were identical to two they used for which ranks given to the pictorial scenes regarding degree of asymmetry by four independent judges were reported]. The scene with the greater degree of asymmetry will be called Scene I, the other, Scene II.

Four by four inch and 8 x 8 inch prints of each of the scenes and their mirror images were used.

For all trials, the 4 x 4 in. print was placed a fixed metric distance of 29 in. from the eye. The larger prints were placed in a carriage which was mounted on a calibrated track. The carriage was manipulatable by *O* by means of a crank. Before each observation, *E* positioned the larger print at a randomly chosen distance along the track. This apparatus and the nature of the room used for experimentation are described at length elsewhere (2).

---

* Received in the Editorial Office on March 14, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

For each scene, I and II, two portions of the scene were specified to *O* as being those which were to be scrutinized during the course of the experiment. Since one designated portion in each scene was in the foreground and the other in the background of the scene, these portions will be called *F* and *B* for foreground and background.

O's task was, on every trial, to adjust the metric distance of the large print so that one of the two specified portions of the scene appeared to be equidistant to the corresponding portion of the scene in the small print. *O*s were instructed to look at only the part of the scene specified, and to avoid looking at all other parts of the scene. Observations were monocular. When the specified portion of a scene was to the left of center, the scene was labelled *l* for the small print, and *L* for the large print; similarly for portions to the right of center for which the labels *r* and *R* were given to the small and large prints, respectively.

Since there were four prints of a given scene, specifying one portion of that scene provided four possible comparisons or pairs of small and large prints, these being *l-L*, *l-R*, *r-L*, and *r-R*. When the designated portion is *F*, these comparisons will be called F(*l-L*), F(*l-R*), etc. With eight different conditions possible for each scene, a total of 16 were possible for the two scenes. Each of the 10 *O*s who served in this experiment made six observations to each of the 16 conditions, for a total of 96 observations. All *O*s met a visual acuity criterion of 20/30.

Five *O*s received the eight *F*-comparisons first, followed by the *B*-comparisons; the other five *O*s received the *F*- and *B*-comparisons in reverse order. Both the eight *F*- and the eight *B*-comparisons were presented to *O*s in prearranged random order. All observations were made in one session, with a 5-minute rest interval between the *F*- and *B*-comparisons.

C. RESULTS

To test the statistical significance of the hypothesis that mean metric distances of the large prints (from the eye) are greater for the *F*-comparisons than for the *B*-comparisons, the data were treated in the following manner. Means were computed for the 160 units of data (16 units for 10 *O*s), each comprised of six observations. The 80 *F*-units of data were compared, singly, with each of the corresponding 80 *B*-units. For example, the *F*-unit of *r-R* for Scene I was compared with the *B*-unit, *r-R*, for Scene I. A "plus" tally was made when an *F*-unit mean was greater than the corresponding *B*-unit mean; a "minus" tally was made for the reverse situation. Assuming that, by chance, *F* would exceed *B* in 50 per cent of the *F-B* comparisons (or 40 comparisons), a test of significance of proportions was made.

In 69 of the 80 *F-B* comparisons, "plus" tallies were made, which resulted in $z = 6.36$, $p < .001$. Therefore, the mean metric distances for the large prints were significantly greater when the foreground objects in the large print were adjusted by *O* to be of the same distance as the foreground objects in the small print than when the background objects in the two prints were matched for distance, the result which was predicted by the experimental hypothesis.

The data were further examined to determine the relationship of locus of objects to be matched (whether in the foreground or background) with the lateral orientation of those objects. Presented in Table 1 are the mean metric distances of the large prints for all *O*s for the 16 conditions. These results are presented graphically in Figure 1. The *F*-conditions correspond to the conditions of the Adair and Bartley study (1), for, though in their study *O*s were told to consider the scene as a whole, *O*s undoubtedly matched the prints largely on the basis of the foreground or dominant objects, since this is an easier and more customary task than matching background objects. As would be expected, then, the progressive increase in metric distance of the large print from F(*l-R*) - F(*l-L*) - F(*r-R*) - F(*r-L*) obtained in this study corresponds very closely to the results obtained by Adair and Bartley. (The metric distances differ in absolute size for the two studies because Adair and Bartley placed the small print 14 in. from the eye; in the present study this distance was 29 in.)

TABLE 1
MEAN METRIC DISTANCES OF LARGE PRINT FOR 10 *O*s FOR 16 EXPERIMENTAL CONDITIONS

Scene	l-R	l-L	r-R	r-L	
			F-conditions		
I	120.80	126.23	129.37	134.36	
II	106.41	113.86	114.67	122.68	
Means	113.60	120.04	122.02	128.52	
					Mean of all F-conditions 121.04
			B-conditions		
I	89.67	91.39	86.50	91.33	
II	94.10	99.09	94.72	98.28	
Means	91.68	95.24	90.61	94.80	
					Mean of all B-conditions 93.14

Mean Scene I = 108.71
Scene II = 105.47
Grand Mean = 107.09

When the results for the *B*-conditions are considered, it can be seen that a

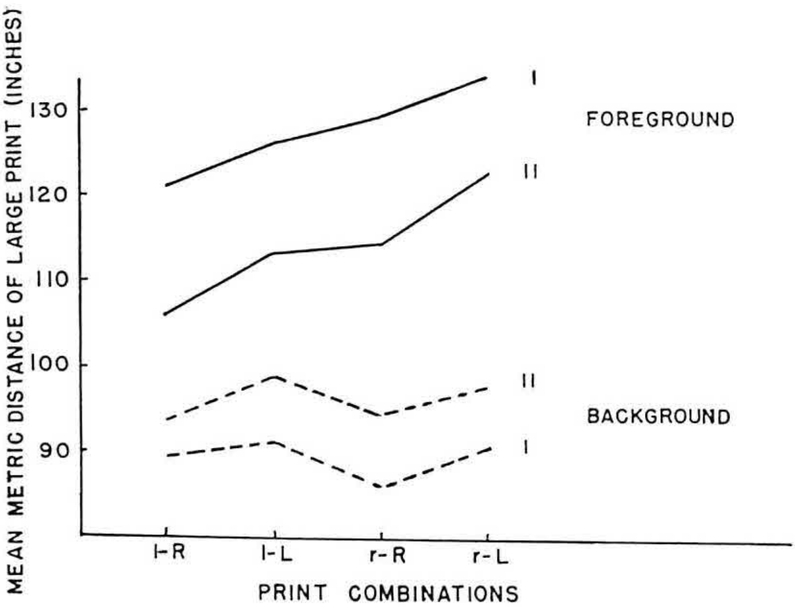

FIGURE 1
The relationship of metric distances of large print for the various combinations of standard and comparison prints. The solid lines are for the crucial items when in the foreground, and the broken lines for the crucial items when in the background. The Roman Numeral I is for one scene and II is for the other. Lower case letters in the abscissa indicate the positions of the crucial item in the small print and the capitals indicate the positions of the crucial item in the large print.

similar progressive increase in mean distance for the large print as a function of lateral orientation of the items in the prints was not demonstrated. Consequently, the relation cannot be generalized to *all* conditions of nearness as a function of lateral orientation in pictures.

For the *F*-conditions, the small print serves as the reference and the lateral orientation of the focal item is crucial to the judgments made by *O*s. This function of the small print appears to break down when distant or background objects are considered by *O*.

Consequently, the relation between lateral orientation of print item and its perceived distance found for foreground items in this study and the perceived distances found in former studies (1, 2) cannot be generalized to cover all print features. When the focally-attended-to print item is in the background in the scene (*B*-condition, i.e., seen as a distant object), the curves of relationship just alluded to do not hold. The conditions of B(*l-R*) and B(*r-R*) resulted in smaller and very similar distances of the larger print than B(*l-L*) and B(*r-L*), which indicates that *O*s' judgments for these *B*-comparisons are dependent on the lateral orientation of the large print rather than upon both prints.

As can be seen in Figure 1, the mean distances of the large print are greater for Scene I for *F*-comparisons and greater for Scene II for *B*-comparisons, owing, no doubt, to the difference in degree of asymmetry for the two scenes.

Furthermore, the *F* and *B* items were not equally distant as far as the subjects' interpretation (perception) of the items were concerned. In Scene I, the *B* item was, by nature of the scene itself, much further away perceptually than was the *B* item in Scene II. The foreground items were not perceptually too greatly different in distance. Therefore, although we call the distant item *B* in both prints, they are obviously at very different perceived distances and ought to make a difference in the outcome. The *B* item in Scene II is approaching what one might call a mid-ground item. Actually, in neither scene was the *B* item on the horizon. This means that despite the fact that the *B* items were not the most distant possible items, they functioned differently than when the items were definitely *F* items.

D. Summary

The foregoing material describes an investigation of whether items in the background of photographic prints look nearer when on the left than when on the right. Adair and Bartley (1), and Bartley and Adair (2) found that items in the foreground followed this rule. The present investigation showed that background items did not follow the rule. The matching technique used to study the distance effects involved matching a small print at a fixed metric distance with a large print at a variable and greater metric distance. In the present investigation, the background item in the small print seemed to matter little, if at all, whether it was in the left or in the right side of the print. This was not the case in the earlier studies of foreground items.

References

1. Adair, H., & Bartley, S. H. Nearness as a function of lateral orientation in pictures. *Percept. & Mot. Skills*, 1958, **8**, 135-141.
2. Bartley, S. H., & Adair, H. Comparisons of phenomenal distance in photographs of various sizes. *J. of Psychol.*, 1959, **47**, 289-295.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 59-74.

THE RETURN OF THE HOMUNCULUS*

Department of Psychology, Michigan State University

LEO L. GLADIN

A. INTRODUCTION

The solipsist as defined by philosophers is a sophisticated creature; he uses the common language to designate the universe of objects and events and conscious beings, after which he denies existence to the whole—thus reducing the field of reality to himself and his subjective experiences. Among philosophers, it is argued that the solipsistic position cannot be disproved, ". . . and is the only *logical* conclusion that can be drawn from the reduction of so called external and objective reality to terms of consciousness and experience" (2, p. 530).

As far as can be ascertained, psychologists have never been cast into despair by this conclusion—logic notwithstanding—and have accepted the external and objective reality as a basic working premise. Nevertheless, solipsism and the more liberal idealistic positions within philosophy do constitute psychological problems whose resolution may well serve to delimit the infrahuman sphere of behavioral operations more sharply, and as well serve to clarify such concepts as "self," "ego," and "consciousness" as they are applied to humans.

B. THE INFRAHUMAN: IMPERFECT SOLIPSISM

The primal organism apparently intruded upon the world-scene without special preparation and without external guidance and control—the realization of which circumstance has caused the existentialist much of his nausea. Chancing to survive and mutate over time into myriad forms, the organism remained isolated from the data of the external universe. During its phylogenesis, however, it had to "learn" to cope with the objects and events of the external field of reality of which it could gain no objective knowledge. Failing to cope with sufficient effectiveness, it was extinguished as individual and often as species. Propitious mutation enhanced its prospects of survival; encumbering mutation reduced such prospects. It did not become modified in structure and function *in order to* survive, since it could neither comprehend the implications of non-survival nor control its own alteration. Obviously,

---

* Received in the Editorial Office on March 18, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the species-representatives upon the contemporary field stem from an unbroken continuity of survivors; by and large, they are a select organismic community—having been molded in the crucible of reality according to its inexorable specifications.

As a late arrival upon the world-scene, and one who has undergone many mutations during the species-span of, say, the cockroach, the human is prone to regard the evolutionary process as at least quasi-purposive—with mutation tending in the direction of producing ever more efficient survivors. However, parsimony would argue for the view that the fact of survival posits neither necessity nor design; it may be accomplished with absurd inefficiency as well as with remarkable efficiency, as the varieties of organismic species will testify. All that is required of the survivor is its presence upon the contemporary field, where it is available for the inspection of the human witness. By an evolutionary trial-and-error process, each variety of organism has somehow solved the problem of coping with the exigencies of its particular portion of the field of reality—none of whose actual specifications have become known to its kind as a result of its "problem-solving."

Paraphrasing an idealistic biologist-philosopher, Cassirer says:

> Every organism is, so to speak, a monadic being. It has a world of its own because it has an experience of its own. The phenomena that we find in the life of a certain biological species are not transferable to any other species. The experiences—and therefore the realities—of two different organisms are incommensurable with one another. In the world of the fly, says Uexküll, we find only 'fly things'; in the world of the sea urchin we find only 'sea urchin things' (1, p. 23).

As a "monadic being," the organism is solipsistic; it is the center of its own special universe which ceases to exist when it expires. It constitutes an organismic point located in physical space whose dimensions are radii extending outward from its psychological center, which assigns significance only to such aspects of its portion of the field of reality as have a bearing on its prospects for continued well-being, and only insofar as it has the capacity to somehow appreciate such significance. (Thus, the beetle, upon becoming exposed to sunlight when its rock is upturned, may seek refuge in the darkness under the instep of the human who is bent on treading upon it; its tropistic reaction is phyletically based on survival-probabilities associated with darkness *per se*—the discrete nature of the *source* having a significance beyond the beetle's capacity to appreciate.) For the infrahuman, reality consists of the sensation-contingent responses it continues to survive; it remains thus a *response-reality* organized around putative objects and events which can do something *to, for,* or *with* the respondent. The object-world remains dependent upon the organism, and "exists" only in such aspects as have an appreciable significance for the organism. Its perspective on the world is subjective and inevitably *internal.* Stemming from a continuity of survivors, it cannot appreciate its own history; existing within an infinitely extendable space, it can only appreciate its immediate surround. It is the familiar "solipsist of the present moment."

Nevertheless, the infrahuman is an imperfect solipsist by the philosopher's standard; it lacks the capacity for self-knowledge necessary for solipsistic perfection. At the human level, such a state of affairs is discovered only during early infancy; according to Piaget: "If the (infant) child really knew himself, we should have to maintain that solipsism exists" (4, *Introduction*). Apparently, as the child develops he does gain some knowledge of himself; however, in the very process of self-realization he loses his eligibility for solipsistic status.

C. THE ORDINARY HUMAN: IMPERFECT REALISM

When he acquired symbolic speech, the human transcended the solipsistic limitation common to the remainder of the animal kingdom, and entered into ". . . a new *dimension* of reality," in Cassirer's words (1, p. 24). Whatever laws of behavior govern the infrahuman cannot be extrapolated to encompass the strictly human condition, although they may have a limited, if significant, applicability to the solipsistic phase of infancy.

The infrahuman is a point-bound unit-individual whose private experiences cannot be communicated to any other individual. In the mammalian series, the organism demonstrates an evident capacity for empathic reaction (3) which tends to modify the behavior of one individual via its exposure to the immediately presented experiences of another, but no meaningful content is transmitted from one psychological center to the other. Empathy gains its functional utility from the respondent's capacity to appreciate the *peripheral* manifestations associated with the private experiences of its fellow. Thus, when the monkey hears a cry of pain uttered by a fellow primate, it may react with affect and apprehensive demonstration even when the fellow is not immediately visible. However, the respondent's empathic reaction redintegrates only *its own* subjective experience of pain and fear on the basis of a qualitative similarity between the cry-in-the-distance and the cry it has heard issue from its own throat during events of tissue injury. Should the injured monkey return from the jungle depths, it cannot inform its fellows as to whether it was attacked by a tiger or a panther, whether it fell out of a tree, or sat on a newly irrupted carbuncle.

Empathy enables the respondent to profit indirectly from the past experience of its kind; thus, the fawn instigates a running response when the maternal doe takes flight without gaining any knowledge of the experiential cause of the doe's alarm. The fawn's flight response may thus become associated with *odor-of-lion* and in the future it need have no direct experience of *actual-lion* such as might well reduce its prospects for continued well-being. Even though such peripheral manifestations as *odor, sound, and appearance* serve as partial bases for empathy, they differ primarily in the *direction* of their organization from similar manifestations upon which antipathy reactions are based; the peripheral manifestations of the lion's proximity are elaborated into flight-signals by the prey animal, and convey nothing of the lion's actual intentions.

The infrahuman may be described as having a uni-dimensional relationship with its object-world; its perspective is point-dependent and sense-limited, and may be depicted as a single line drawn between its point-location and any *significant* "object" of its experience. The human, on the other hand, may be described as having a tri-dimensional relationship with any *designatable* object of the external universe. By means of a commonly shared syntactic language form, any pair of communicants comprise the base points of a hypothetical triangle whose apex is the object under their joint consideration. Between them, the pair can demonstrate that the object-as-sensated is not the exclusive experience of either of them and therefore external to both and independently existent. Thereafter, any communicant who has engaged in this inevitable objectification experience is able to realize the existence of objects when he is by himself.

At the same time, he discovers that the human objects of his surround have experiences different than his own, and that he is himself an object to his fellow communicants—whose day-by-day commentary on his appearance and behavior furnish him with an external knowledge of himself. Becoming an object unto himself, he is enabled to realize his own mortality—the inevitability of his own demise—and must also realize that the universe will continue in existence beyond his lifetime, as it has beyond those of his sometime fellows now deceased. Without giving thought to the process, the human community thus pools its individual experiences and verifies the existence of an object-event universe external to—and independent of—the immediate presence of any sensating individual.

For the point-bound individual, any sensory experience of a putative object constitutes a *phenomenon qua appearance;* should an infrahuman mistake a fluttering leaf for a butterfly, it can have no appreciation of its error. Each *appearance* has its own internal validity, and instead of recognizing an error, the infrahuman must assume that one appearance merely supplanted the other. As a communicant, if Peter states, "I see a butterfly," Paul may respond, "No, it is a fluttering leaf." As a witness to Peter's experience and testimony, Paul serves as an external corrective—if his own report is veridical. When Peter now arrives at the location of the fallen leaf, he is enabled to realize the possibility of error as one of his human potentials. Thereafter, even when Peter is alone he is able to suggest to himself that he has "made a mistake" under circumstances appropriate to such a realization.

Since error is potentially correctible, the human's discovery of the possibility of error is a universe of discourse distant from the solipsistic dependence upon appearances. The ordinary human thus obtains access to the actual field of reality; although a great deal of his purported knowledge of the field is contaminated by error, he has transcended the idealistic limitation which binds the most complexly organized of infrahumans. Even when the social consensus of his sub-community subscribes to a gross error in place of a veridical description of a datum of reality (as when a sect insists that the earth is a plane surface) the human's external perspective places him within reach of reality. Since the field of reality does not alter to accommodate an erring social consensus, it never assumes the status of an *appearance* as a consequence of the group's non-veridical testimony. In comparison with the mortal flux, the physical field of reality is relatively stable; any item of knowledge which is congruent with an aspect of the Real will endure beyond the moment-to-moment aberrations of the individual or his consensus group. Being possessed of an error-concept, any knowledgeable individual owns the potential to participate in a correction of the social consensus. Nevertheless, the process is not a rapid one, so that the field of reality as it is viewed by the several sub-communities is variously seen, according to the cultural history of their populations. Thus, the ordinary human has access to the field of reality but may lack the capacity, the knowledge, or the willingness to subscribe to the data of reality.

D. Denotation and Human Motives

The spring blight may destroy next autumn's banana crop, and there may be a current boom in the monkey market stimulated by researches in progress on other continents, yet the primate in its native habitat can have no awareness of such matters; its concerns are sufficient unto the present moment. The price of wheat as quoted in New York may cause a man in Tokyo to leap from the nearest tall building, and at least a few persons experience a touch of nausea at the thought that the sun will burn out some millions of years hence. The extrapolative function of language which enables the human to overleap time and place begins to make its impact on the human when he is born, if not sooner. Since the anxieties of his animate surround affect the human during his period of intellectual and physical incompetence, they play some part in determining the direction and limits of his motivation.

Marks on paper made by men whose earthly remains have disintegrated over the thousands of man-designated years continue to be sources of remarkable and diverse motivation in the several sub-communities. The written "thoughts" whose *thinker* has no other discernible existence may be translated into many other syntactical forms, and afterward may be written, spoken, or spelled into the hand of the receptive human; As Cassirir observes (*op. cit.*), their meanings demand no specific physically energic order of transmission—which fact puts them a universe of discourse distant from the energic orders upon which the infrahuman's access to its special environment must inevitably depend. Burning the inscribed parchment leaves no residue of meaning in its ashes, nor will such meaning be discovered in its smoke in the form of "free energy." As objects of the physical world, syntactical forms are inert and yet have a profoundly significant potential which cannot be described in terms of physical energy.

Although meanings are derivable from the object-event universe via the experiences of an appropriate subject, and by means of denotative symbols are transmissible between subjects, they can have significance only within the psychological center of a human organism, according to his capacity and sophistication. Within the individual, they are infinitely combinable; the pinkness of the possible cloud gives way to the "pinkness" of the improbable elephant. Via language, the human extrapolates fictions fully as motivating as are his observable realities.

The genesis of a communicating society out of a gregarious herd must necessarily have involved a motivational expenditure directed toward an adequate denotation of the object-world. Beyond the immediate utility of objects, the human was required to observe them for purposes of denotation—discriminating the differences out of similar things and discovering the similarities among different things in order to establish the reliability of communication. Whereas the central nervous system of the infrahuman simply neutralizes most of its environmental effects into a background for its significant experience of the world, the human could allow no such vast lacunae once he began to designate the objects and events of the universe. This process in itself must have required a concentration of attention upon details of the environment which could have no possible significance to the most complexly organized infrahuman. By a simultaneous process of denotational elaboration and sharing of formerly private experiences, the human objectified the universe and was at the same time exposed *to* its objects and events, which he was motivated to investigate because they were confronting him, and not because they could do something for, to, or with him. Whereas he previously responded only to the significance which a relatively few objects and events had upon his continued well-being, the emergent human's objectification rendered all things potentially significant to the continuity of humankind.

E. THE HUMAN'S EGO STRUCTURE: HOMUNCULUS?

The philosophical plausibility of the solipsistic position lies in the tacit assumption that a knowledgeable *human* organism is a unit-individual whose denotational specifications require no reference to a community of similar organisms. However, if the human qua human has a "self," "ego," and "consciousness" whose emergence is community-dependent, the subject terms of solipsistic propositions become plurals and therefore cannot very well take singular predicates.

The archaic dictionary definitions of consciousness: "mutual awareness," and "awareness of that of which another is aware," suggest that *consciousness* is something other than *awareness*. An infrahuman may be aware of a sensation of pain or of the presence of an object in its visual field (as can be inferred from its immediate behavior) but it cannot be aware of the possibility of fatal infection from the wound which causes its pain, nor can it be aware that once the object was removed from its visual field it was carried half way around the world by an aeroplane. In order to have this level of awareness, it would have to transcend its unit-individual status by participating in a communicant relationship with others (*conscious:* "sharing knowledge").

The human, on the other hand, cannot be aware (in terms of sensation) of what his liver is doing at any given moment, but he can be conscious of what, in general, it ought to be doing. His knowledge is gained from the study of other livers than his own by persons skilled in specific investigatory techniques. For the infrahuman, "other livers" can only carry a *to-be-eaten* significance, never a functional one. The human can be conscious of having been invaded by minute organisms when he suffers tissue injury, and may be motivated to apply iodine-and-bandage in order to avoid a more serious, if not fatal, consequence. Yet his senses cannot make him aware of the presence of invading organisms or of the implications of fatal consequence; he can envision himself moribund, deceased, mourned, buried, and forgotten or cherished in memory (but not in *his own* memory) and thus reveals that he has transcended the physical dimensions of time and place contingent upon his physiological mortality.

Once he has gained a consciousness of the proportions and implications of the object-event universe, the human does not require the presence of a fellow communicant in order to operate effectively upon the field of reality; he may reflect, judge, compare, and discover on the basis of a "mutual awareness" which has become a covert and private function. He is enabled to investigate the relationships among objects and events *apart from* himself whereas the infrahuman can only evidence an interest in putative objects and events as they are somehow related *to* itself.

The occasional discovery of feral children, who have presumably spent their developmental years outside of a communicant society, offers an exemplification of apparent human form in the absence of human function. While such a deficiency is sometimes explained by the inference that ferals were mentally incompetent prior to their isolated existence, the "explanation" itself presents the interesting proposition that a mental defective is able to shift for himself within a primitive habitat whereas he ordinarily requires special protection during his developmental years within a cultural surround. As a human, Helen Keller was grossly defective until the age of seven; then, deaf, dumb, blind, she discovered the principle of symbolism, "with its universality, validity, and general applicability . . . giving access to the specifically human world, to the world of human culture" (1, p. 35).

Whereas the philosopher may describe consciousness as the spiritual essence which pervades the universe, the psychologist must operate with greater constraint; his several versions of "consciousness" are conceptualized as organismic properties and often confused with simple states of *awareness*. Since consciousness is a covert function, at least some psychologists prefer to regard it as ineffable and therefore not a suitable subject for serious scientific consideration; however, if consciousness can be described in terms of symbolic operations conducted by organisms having certain specifications, it merits attention as one of the most significant developments, if not the most significant, of the organismic sphere.

As a property of the human organism, consciousness may be conceptualized as an attribute of *ego*. Whereas the infrahuman has a "brute self"—a neurophysiological point-bound structure which organizes its awarenesses of its actions upon and impacts from its subsistent environment, it cannot be said to have an "ego." It cannot locate its "self" as an object among objects and base its actions upon such a transcendent perspective. In fact, it cannot be said to act upon itself at all, since this would constitute an illogical tautology. It lacks that double capacity of subject and object which is so apparent in the human.

Whereas the philosopher may state the specifications of a "pure" ego, which can act without reference to organismic frailties, the psychologist must locate his conceptualized ego upon a more palpable base within the neurophysiology of a human organism—where it is liable to both gross and subtle contaminations. As a member of the continuity of organisms, the human is a material point located in physical space, and he has certain material requirements in common with other organisms; in addition to ego, he has a brute self with its propensities for organizing a sphere of operations in accordance with the laws of behavior applicable to infrahumans. The functions of empathy and antipathy, as well as of primitive identification and imitation (3) gain their significance for the respondent as a result of its experiences of the *peripheral manifestations* of the animate object; as such, they represent signal functions which enable the respondent to establish directional affiliations with its animate environment according to the propensities laid down during its phylogenetic experience. Of such signal functions, Cassirer remarks: "Signals and symbols belong to two different universes of discourse: a signal is a part of the physical world of being; a symbol is a part of the human world of meaning. Signals are 'operators'; symbols are 'designators.' Signals, even when understood and used as such, have nevertheless a sort of physical or substantial being; symbols have only a functional value" (1, p. 32).

The human infant, which begins its life in a condition of imperfect solipsism, makes its initial affiliations with the animate surround on a signal basis, since it is not otherwise equipped to become related to it. Lacking an ego structure, its self-world relationships are elaborated with itself as the center of its universe, so that its affiliations with putative objects are accomplished by operatively (on an involuntary conditioning basis) ascribing to them significances determined by the activation of affective propensities. Thus, the aroused affect structures the putative external situation according to the significance of the peripheral manifestations of the animate surround; the actual intentions of the object are unknown and unknowable. The simultaneous presentation of a loud sound and a white rat—which had no *actual* relationship with one another—became related within Watson's infant subject, Albert. The affect aroused by the sound structured the rat as a "fear-thing." Such affective conditioning can become strongly organized, causing the respondent to regress to the infrahuman mode of reaction even when he becomes a knowledgeable adult. Thus, the human who fears mice may know that he is stronger than they and that their inferable intentions are peaceable —*as long as he is not in the actual presence of a mouse;* yet the appearance and movements of an immediately present mouse will rouse within him an affect of fear which furnishes the animal with the significance, "wild-beast-intent-upon-attacking-me," even though the disinterested observer notes that the mouse is seemingly panic-stricken and in full retreat. Such a pattern of behavior bears more resemblance to *reversion* than to regression in that the objective knowledge upon which ego-control of behavior is founded loses its functional effectiveness as the respondent returns to a subject-dependent affiliation with a putative object whose significance is affectively defined according to an internal perspective which cannot take into account the *actual* specifications of the object-in-itself. This mode of behavior, which is common among infrahumans, effectively discourages "reality-testing"—the alteration of responses toward congruence with the actual situation as it exists from moment to moment. For the infrahuman, it is the *sufficient* mode as long as it offers the solution of an immediate problem of contact with the object-event world.

Since the human infant enters into life as a solipsistic organism, and is physically and mentally incompetent to cope upon the field of reality for many of his developmental years, signal processes may enlarge his "brute-self" structure while he is gaining a mastery of symbolic functions which enable a consciousness of himself and of an external universe of denotable objects and events. While his *ego* "consciousness of himself" is becoming defined as an object among objects which must accommodate itself to the social realities of human individuals having specifications unique unto themselves, his *brute-self* structure may be elaborating the human surround according to signal processes which classify grossly according to affectively significant peripheral manifestations.

By virtue of their being *conditioning* operations, signal processes are not available to the respondent as items of ontogenetic acquisition—unlike his formal and informal learning of symbol functions. Nevertheless, he begins to *operate* according to their dictates during his preverbal phase of affiliation with his animate and inanimate surround, and the behavior patterns so established are undoubtedly reinforced aperiodically throughout his lifetime according to the infrahuman mode—which effectively prevents him from gaining an accurate and utilizable knowledge of other humans having specifications unique unto themselves. Empathy-non-empathy and empathy vs. projection studies offer supportive evidence for the above-stated point of view

(5). The "unconscious fantasies" reported by Freud and others may represent the analysand's early experiences of affective signal-process affiliation with his animate surround.

Because the human very often develops within a psychopathological surround, he is forced to somehow adjust to unreality as he develops as well as to the actual field of reality. Since pseudo-knowledge can be acquired by the respondent in the same manner as he gains knowledge of reality qua reality, and has a functional validity by social consensus, it enters into the economy of the ego structure. Nutritional, libidinal, and evacuatory functions may thus be viewed according to a distorted perspective which prevents the smooth transaction of inevitable and common physiological operations. The human is put "at odds with himself" in the same manner as when he reverts to the infrahuman mode in rage and fear reactions and is afterward "ashamed of himself" for behavioral displays contrary to the social code which offers maximum benefits over the long run.

On the other hand, societal pseudo-knowledge can be entirely harmonious with the infrahuman mode of reaction in any given sub-community, and the symbol system may be employed for its reinforcement. Thus, the peripheral manifestations which form the fundaments for such signal processes as *prejudice* and *superstition* can provide the sufficient mode for categorization of persons and events. The several respondents react to a feeling which "recreates" the object as a "fear-thing" or "anger-thing," in the case of prejudice, with instances of positive correlation between affective delineation and behavior of the object being confirmatory and negative instances being unobtainable when the respondents are in the presence of the object due to the suspension of empathic capacity in favor of antipathy. *Theoretical* knowledge of the object of prejudice may be obtained, but of itself must lack the validity derivable from organic responses of a positive nature, on which the functional effectiveness of such knowledge depends. In both superstition and prejudice, confirmatory instances can be readily obtained via verbal communications, so that the beliefs are reinforced in the absence of individual experience. If Henry dies and Peter has not heard a dog howl prior to his passing, it is sufficient that Paul received the omen and reported it.

Since the ego organization may be contaminated in its representation of both the internal realities of physiological function and the external realities of objects and events, the ordinary human ego must be regarded as "impure"; nevertheless, even if it is conscious of spurious orders-of-things it remains the representative of the external world, and must somehow arrange a compromise between the more or less distorted impulses of its brute self and that world. Its strength and control depend upon physiological energy, but the extent that it gains functionally effective knowledge of reality probably depends more on the degree to which the brute self has escaped the distorting effects of early affective impacts than any effort of "will power." Given a basically sound physiological structure the human has the potential capacity to accept pain, nutritional and libidinal deprivation, fatigue and community disapproval in pursuit of goals from which he can hope for no gratification in his own lifetime—thus going beyond anything possible in the animal kingdom of the infrahuman, who lacks the perspective to see beyond the sensation-dictated moment into all time, all place.

Although the Cartesian notion of *homunculus* is one of the prime anathemas of general psychology for the reason that "it doesn't explain anything," and in addition posits a little man inside the big man, the common speech, as well as scientific exposition in the realm of ego theory, employs sentences which describe activities in terms of subject-object relations within a single individual. Such sentences ought to constitute tautologies if the human is a unit individual whose "self" is indivisible, conceptually or otherwise. If, on the other hand, the human has *two* selves whose organizations are laid down along different lines and whose perspectives are mutually contradictory, the tautology does not exist. The contriety which is expressed in "man against himself," is expressed by the layman in such sentences as, "I was angry with myself," "I was ashamed of myself," "I have to drive myself," and "I must get control of myself."

Other sentences regard the "self" from the outside, as a designated point in physical space: "I did it for myself," "I will defend myself," "I bought it for myself." Here, the apparent tautology recognizes the existence of other humans; if one did not recognize himself as an object among objects, and was therefore a true solipsist, it would be redundant to specify *himself* as the object of any predicated action. There would be no denotable *other* to "do for," "buy for," or "defend."

Where the human expresses a contriety of feeling about himself, he generally refers to behavior or ideation which represents either brute impulse activity which has been communally defined in negative terms or normal physiological impulse activity which has been misconstrued in negative terms by his consensus group. This group may be as small as his biological family unit, yet it is likely to be the determinant of his ego perspective upon himself in many areas of interpersonal contact because it makes its impact during the period of his physical and intellectual incompetence. This familial impact results in a "superego formation" which is probably neither inevitable nor necessary—an erroneous prototype of social reality which is inflicted upon an extremely labile, extremely defenseless, central nervous system. Whether his perspective on himself is accurate or not, it is nevertheless founded in his externalized, knowledgeable ego. If his ego has been invaded by facets of the brute mode, such as prejudice and superstition, he need not be ashamed of impulses so based if his sub-community is not; if his knowledge is contaminated by fictions to which his sub-community subscribes, he may be ashamed contrary to reality, and if not ashamed perhaps defiant, anxious, or apathetic. In either instance, he views himself "from the outside" as a result of his community experience in becoming an object among social objects.

Freud's discovery that the analysand *resists* self-revelation points to a situation which is more real than apparent; it is not the analyst to whom he primarily makes his resistance but to himself as a communally defined, communally maintained being—many of whose private experiences (both actual and ideational) are currently reproachable and possibly punishable both here and *hereafter* and have in the past been actually punished or "punished" in the sense of accidental impacts occurring contiguously with act or impulse.

However, the reason why he cannot reveal many of his significant experiences despite skilled, patient assistance from an accepting, non-judging, sympathetic professional probably lies in the fact that the significance of his every experience is internally contained. If objectified and externalized, his universe *as he experiences it* is nevertheless within him and represented in his neurophysiology. Whether he is participating in community activity or is alone, the community continues to be represented within him in terms of its impacts upon him and his transactions with its objects and events. It is the very fact that the community is functionally effective within him when he is in separation from his social others which makes it possible for him to behave in the human mode at all; the symbolic process which enables him to discriminate the several aspects of a bland object enables him to suffer pangs of "conscience" over a private bit of ideation. Although he *knows* he is alone (having cognitively structured "aloneness" as something having different specifications than community participation) his necessarily internal organization of the field of reality may bear heavily upon him for having proscribed thoughts whereas it offers no negative provision against "inspecting-bland-objects." Since it is not the social reality as such, but the internal organization of its significance which modifies and controls his behavior, he may be overly inhibited or wildly impulsive, may defy the social conventions or attend upon them with such scrupulosity as to afford himself very little gratification. Even if nothing in his current surround can injure him, he may learn to believe, against reality, that he is in danger of annihilation by objects and events to whom he has erroneously given their definition, their *meaning*. Quite obviously, the meaning does not reside in the object or event of itself, since he can discover certain of his fellows disporting themselves among such without harm (save such harm as he may *wish* upon them for behaving as he cannot).

It is the psychopathology made possible by syntactic forms which most clearly shows that the human as a social organism is a homunculus—a synthetic *man* who has transcended a point-bound, sense-limited status by including the human community within himself—enabling him to become more-than, less-than any other organism within the animal kingdom.

F. SCIENCE, REALITY, AND FIELD-PERSPECTIVE: THE PURE EGO

As scientist, the human abrogates his "self" and becomes (at least in his professional capacity, and often only theoretically) a "pure ego." His operating perspective is assumed from no necessarily specific point-location, is not sense-limited or hampered, and is not consensus-bound. It observes no national boundaries, and stretches across time beyond ordinary comprehension and space so vast that it can only be described and handled mathematically in terms of velocity-per-time.

The scientist obtains his perspective from the field of reality itself, according to his accumulated knowledge of its myriad aspects; if he does not directly observe very much of what his mind "sees" (as is the situation of the atomic physicist) at least he sees to it that no item of purported knowledge is firmly held if it is somehow incongruent with the field of reality; even when an item of knowledge *has* congruence, it is held with some degree of reservation, since it is realistic for him to assume that he does not know everything about anything. Observing an electron-path in a cloud chamber, on a photographic plate or other adjunct to his relatively myopic human vision, he demands of his ego that it make his theory congruent with his fleeting glimpse of the infinitisimal; where one observation contradicts another, he is proscribed from a denunciation of fickle reality and searches instead into human fallibility, of whose dimensions he has become conscious over the arduous years of discovery of things-as-they-are out of things-as-they-seem.

It is in the coherent fitting of knowledge to the data of reality without a seam of contradiction that his hope of ultimate realization lies—wherein the order of past knowledge must inevitably yield future knowledge, untainted by his private fears, his hopes, his heart's desire. This is his significant, most sufficient motive—it constrains him from treasuring his items of contribution to the fund of knowledge as personal possessions to be handled with care, like the delicate children of a fond parent; his brain-children spring forth full-armored, in need of no defense, to stand and be measured against the field of reality—the majesty of all time, all place.

G. SUMMARY

It was proposed at the outset that solipsism and the more liberal idealistic positions within philosophy constitute psychological problems whose resolution may serve to delimit the human from the infrahuman spheres of operation more sharply, and also serve to clarify such concepts as "ego," "self," and "consciousness."

The infrahuman was described as a sense-limited unit-individual—the nature of whose environment must continue dependent upon its location as a phyletically structured point in physical space. It is thus the center of its own private "universe," and it can neither know anything of the experiences of other such centers nor of an external and independent object-event universe which has a spatially and temporally extended existence beyond its mortal time and place. Its perspective on its surround is internal, being limited and defined according to its own inevitably private experience of the external world. The infrahuman is thus an imperfect solipsist which can gain no knowledge of itself as an object among objects.

Since by means of his syntactical symbol systems of communication the human is enabled to share the private experiences of his fellows, he can gain knowledge of an existent external universe and of himself as one human object among many. By sharing knowledge, he becomes conscious of himself and of an independently existing universe. His communally defined "self" as *ego* is enabled to execute its functions according to a perspective upon the actual field of reality: an external perspective which allows him to transcend the limits of sense and his immediate point-location. However, his ego function is liable to contamination because it is located within a physiological organism whose larger history of response relations was evolved during his infrahuman phylogenesis.

As a human qua human, his ego structure constitutes a homunculus synthesized out of the individual's participation in the experiences of the human community; as *homunculus,* his ego is enabled to act when he is in separation from his fellow communicants, even though its organization and maintainance is community-dependent. His universe of *meaning* is contained within him, and motivates him according to his own private organization of the significance of all external orders-of-things. His ego is therefore only as pure as his knowledge of reality is accurate.

As scientist, the human is enabled to have an ego function which is "pure" in that his "self" and its frailties are abrogated according to procedural rules designed to the end of the attainment of knowledge of the field of reality as it exists "of itself." His scientific perspective is obtained from the field of reality via his current knowledge of its specifications—with which his further knowledge must be made congruent if it is to have the status of validity.

REFERENCES

1. CASSIRIR, E. An Essay on Man. New Haven: Yale Univ. Press, 1944.
2. FULLER, B. A. G. History of Philosophy. New York: Holt, 1945.
3. HEBB, D. O., & THOMPSON, W. R. The social significance of animal studies. In: *Handbook of Social Psychology* (Lindzey, G., *Ed.*). Cambridge: Addison-Wesley, 1954.
4. PIAGET, J. The Construction of Reality in the Child. New York: Basic Books, 1954.
5. SARBIN, T. R. Role theory. In: *Handbook of Social Psychology* (Lindzey, G., *Ed.*). Cambridge: Addison-Wesley, 1954.

General Delivery
West Side Station
Colorado Springs, Colorado

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 75-84.

THE MINORITY NEWCOMER IN OPEN AND CLOSED GROUPS*

Fels Group Dynamics Center, University of Delaware

Robert C. Ziller, Richard D. Behringer, and Jacqueline D. Goodchilds[1]

A. Introduction and Problem

The degree and rate of assimilation of a newcomer to the group may be analyzed in relation to such components of the total situation as the perceived characteristics of the host and donor groups, the characteristics of the newcomer as perceived by the host group members, the reason given by the newcomer for shifting group membership, the introduction procedures, the relationship between the host and donor groups in the more inclusive environment, and the rate of change of members from the donor to the host groups. In the present study of the assimilation of the newcomer, two of these general conditions were varied simultaneously: the nature of the host group (open or closed) and the perceived relative status of the donor group (Negro or White).

The open-closed group concept derives rather broadly from a general systems theory framework. A basic unit of this approach is a "system" defined as a "set of elements standing in interaction" (2, p. 23). Systems, in this very general sense, may be categorized as either "open" or "closed." A closed system is considered to be isolated from its environment: new elements are excluded and the departure of elements is precluded. However, it is proposed that most living organisms are essentially open systems since they are maintained "in a continuous inflow and outflow, building up and breaking down of components: never being, so long as it is alive, in a state of . . . equilibrium but maintained in a so-called steady state . . ." (2, p. 23).

Extending this framework to social systems, an open group may be defined as an interacting set of persons in a continuous state of membership flux. On the other hand, membership is static in closed groups.

In an earlier experiment (7), it was hypothesized, generally that a new mem-

---

* Received in the Editorial Office on March 21, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was part of a research project under Contract NONR-2285(02) with the Group Psychology Branch, Office of Naval Research. This paper constitutes Technical Report No. 1 under that contract.

ber is assimilated with greater facility by open groups as compared with closed groups. Underlying this proposal was the assumption that members of open groups in contrast to closed groups perceive idiosyncratic inter-dependence as dysfunctional. Under the open conditions or conditions of social metabolism where the elements of a complex are temporarily inter-related (due to the anticipated replacement, removal, or addition of elements or members), constant adjustment to the peculiar characteristics of the changing membership tends to dissipate the energy of the group in its endless efforts to achieve even a modicum of group structure.

Group syntality (4), the measured performance of the group as a whole, offers a more durable basis for group structure and meaning. Thus, open groups are presumed to be more task oriented; closed groups more interpersonally oriented. Accordingly, it may be anticipated that open groups will be less concerned with who actually performs the task. The mere fact that the task is accomplished is a sufficient condition for member satisfaction (3, p. 310). This would also suggest that the newcomer as a stimulus or press in Murray's terms (5) is more likely to be perceived by open groups as compared with closed groups as a potential group resource with reference to group productivity rather than as a restrictive and/or disruptive supernumerary with reference to existing patterns of social cathexis and interdependent task functions. Or perhaps more simply, the newcomer enters a far more complex, interrelated system of elements in the closed groups than in the open groups, and assimilation is accordingly accomplished with reduced facility and increased hostility—at least during the initial phases of the assimilation process.

The results of the earlier experiment generally supported the aforementioned framework. The newcomer was assimilated with greater facility by open than by closed groups. However, the present study was designed to investigate the assimilation of the newcomer when the open-closed group dimension was varied simultaneously with the second variable under consideration, the nature of the donor group as represented by the newcomer.

In an earlier field study of the assimilation process by Shuval in Jerusalem (6) it was concluded that the crucial factor "conditioning the choice of targets of hostility and affinity" appeared to be the predominant class membership of the ethnic group. Furthermore Shuval states that "as long as ethnic membership provides a highly visible criterion of differentiation, stratification is likely to occur in terms of this criterion." The present experiment was designed to test these conclusions in a laboratory setting involving Negro and White ethnic donor groups. However, in view of the earlier considerations concerning the relative salience of interpersonal orientation and interdependent adjustment in open and closed groups, it was also anticipated that the differential acceptance of minority versus majority newcomers would be less pronounced under the open as opposed to the closed group conditions.

B. METHOD

1. *Subjects*

The 64 freshman ROTC students comprising the subject pool were assembled at random into 24 three-person groups, each containing two naive subjects and one confederate.

2. *Procedure*

The experiment was described to the subjects as a study of group decision-making behavior. The host, three-person groups (two naive subjects and a confederate) were required initially to solve a Meier Art Judgment task through the medium of a standard note-passing apparatus.[2] Prior to this task however, the "open" groups were informed that sometime during the experiment a member would be removed from the group, or a new member would be added to the group, or a member would be replaced by a new member. In addition to the foreknowledge of impending membership changes, under the open conditions one of the initial three members (a confederate) was replaced by a second confederate just after the group had completed the first task. Under the closed conditions, the original confederate remained in the group throughout and the group was never informed of the impending membership changes.

Following the completion of the second task (a relatively simple dot-estimate problem) a new member, either Negro or White, was introduced to the group on the pretext of studying group size effects relative to group decision-making behavior. Finally, the group was presented with a second dot-estimate task but one which was considerably more difficult than the first. The newcomer in each case was, indeed, another confederate who knew the correct answer and a correct method for arriving at the answer. (Thus at this point, the group was composed of two naive subjects and two confederates, the newcomer and the member in the group from the outside or his replacement.) The criteria of assimilation included the degree to which the host group was persuaded to accept the correct answer, the member's degree of estimate change, self reports of estimate change, three ratings of the newcomer by the naive host group members, and the number of notes to the newcomer by the regular members.

---

[2] Essentially the note passing apparatus consisted of a circular table with panels precluding visual communication among members. A communication net block in the center of the table contained slots through which every member was able to send notes to every other member.

a. The critical task. The third problem and the critical task (from the standpoint of the experimenter) was exceedingly more complex than the previous dot-estimate task since the figure contained about 10 times as many dots and these were presented in a form resembling a ping-pong paddle rather than a rectangle. The slide containing the figure was displayed for only five seconds. Following this display the members were asked to submit individual judgments as to the correct number of dots on the slide. When the members had submitted their appraisals the group was informed that they would have "15 minutes in which to reach a group decision as to the correct number of dots on the slide." Finally, after the group estimate had been submitted, the individual members were required to submit a private estimate of the "number of dots they personally really thought there were on the slide."

b. The confederate's role. The confederates received detailed instructions as to their roles. On the bases of previous studies (7, 8) initiating messages and responses were developed. These were used almost exclusively by the confederates throughout the note-passing session.

The initial message sent to all naive members by the newcomer laconically gave the correct "estimate" and a correct arithmetic process for arriving at that estimate. Thereafter, the knowledgeable newcomer merely responded to inquiries with prescribed messages which reflected the attitude that he was highly confident of his estimate but was ready to compromise for the sake of group efficiency and harmony. Moreover, both confederates were enjoined to avoid the "leadership" role, that is, not to assume the function of submitting the final estimate. It should be noted that the task was of such extraordinary complexity that the naive members could neither estimate the correct number of dots with confidence nor logically determine or happen upon the correct approach during the extremely limited time exposure of the slide.

On the other hand, the newcomer focussed on the rectangle of minimum dimensions which enclosed the paddle-shaped dot display. This paddle-shaped figure covered approximately one-half the area within the rectangle. Thus, the number of dots within the figure could be approximated by estimating the width and length of the rectangle in dots and dividing the product by two. The figure contained 1050 dots.

It will be recalled that in order to maintain a constant group size the experimental procedure required the use of two confederates, the newcomer and a regular member who was in the group from the outset or who replaced the first regular member just prior to the first dot-task. The host-group confederate played a role similar to that of the newcomer. However, while he employed a problem-solving process similar to that of the newcomer on the last and critical task, his proposal was a gross underestimate. Moreover, on the earlier tasks, the regular confederates always advocated the correct answer but offered little resistance to group pressures.

c. Measures. The dependent variables were essentially the same as those reported in two previous studies concerning the assimilation of the newcomer (7, 8). These variables include measures of communication direction, group influence, and the relative ratings of the newcomer by the members of the host group. It is presumed that each of these measures reflects some aspect of the degree of assimilation.

The measure of communication direction was simply the total number of messages sent by the naive subjects to the newcomer.

The first measure of group influence was simply the absolute difference between the group decision and the correct answer. The second measure of group influence was derived from the ratio of the degree of estimate change in the direction of the correct answer to the degree of possible change. An upper limit of 1.00 is inherent in the index and a lower limit of zero was imposed. The resulting ratio was subjected to the arcsine transformation. The mean of the members' transformed indices provided the group index of estimate change.

A phenomenological measure of influence was also included. This measure was derived from the mean of the weighted responses to the naive members of the following item: "I changed my individual estimate of the number of dots a great deal as a result of the group discussion." The alternatives were arranged on a six-point scale varying from "agree very much" to "disagree very much."

Finally, the naive members rated themselves as well as the other group members (including the confederate) on nine-point descriptive scales regarding the following statements: (*a*) This person's ideas with reference to the dot-estimate task were of the very highest quality. (*b*) This person is extremely pleasant and contributed a great deal to making this session enjoyable. (*c*) This person exerted a very great influence throughout the meeting.

C. Results

Analysis of variance applied to each of the assimilation measures revealed no statistically reliable effects associated with either of the two condition variables. However, it is apparent from the several means for the four treatments (Table 1) that the closed groups with the Negro newcomer stand apart from the other treatments. Thus, under the closed-Negro conditions groups reached the least accurate group decisions, changed their initial estimates least, and reported the lowest degree of change. The fact that these same groups also showed more individual agreement with the group answer may indicate that the newcomer in this situation had relatively little impact on the group.

TABLE 1
A SUMMARY OF THE MEANS OF THE DEPENDENT VARIABLES UNDER THE FOUR EXPERIMENTAL CONDITIONS

Dependent variables	Experimental conditions			
	Open-White	Open-Negro	Closed-White	Closed-Negro
% Notes to newcomer	64*	68	73	67
Group decision error	350	371	308	595
Estimate change	.61	.70	.56	.45
Self report of estimate change	7.50	6.67	7.67	6.00
Satisfaction with group decision	9.61	10.76	13.46	7.41

* None of the differences are statistically significant.

As a control for the possibly attenuating effect of individual differences in Negro prejudice, the subjects were divided into high and low acceptance categories on the basis of their ratings of the newcomer concerning "quality of ideas" and "pleasantness." Since the Negro and White newcomers sent essentially the same messages, any differences in the ratings of the Negro and White peregrinators may be attributed to differential reactions to physical appearance, that is, ethnic group membership. There was a tendency for ratings on the two separate items to be positively associated ($r = 0.50$, $p < .05$), but it must be noted that the correlation was higher for ratings of the Negro newcomer ($r = 0.66$, $p < .01$) than the White newcomer ($r = 0.35$, ns). The two correlations do not differ significantly however.

Scores on the two individual measures of influence (self report and estimate change) were subjected to analysis of variance in a 2 x 2 factorial design involving the open-closed dimension previously described and the new prejudice classification derived from the summed ratings of "quality of ideas" and "pleasantness" of the newcomer. Division into high or low prejudice classification was at the median and ignored group membership. Cases were removed at random from the resulting four cells to provide an equal number of subjects within each treatment. Analysis was perforce done separately for the White and the Negro newcomer condition. While degree of acceptance of the White newcomer shows no significant relationship to influence (Table 2), the relative acceptance of the Negro is definitely related to the degree one is influenced by him (Table 3). Under the Negro newcomer conditions, the average estimate change ratio (an arcsine transformation of the amount of estimate change divided by the amount of possible change) of accepting subjects was 43.6 as compared with 12.1 for the rejecting subjects.

TABLE 2
ANALYSIS OF VARIANCE OF ESTIMATE CHANGE WITH REFERENCE TO OPEN-CLOSED WHITE GROUP CONDITIONS AND SUMMED RATINGS (HIGH OR LOW) OF THE NEWCOMER

Factor	Mean square	d.f.	F
Open-closed groups	139.92	1	——
High-low ratings	1320.65	1	1.44 (n.s.)
Interaction	890.84	1	——
Between	916.15	16	

TABLE 3
ANALYSIS OF VARIANCE OF ESTIMATE CHANGE WITH REFERENCE TO OPEN-CLOSED NEGRO GROUP CONDITIONS AND SUMMED RATINGS (HIGH OR LOW) OF THE NEWCOMER

Factor	Mean square	d.f.	F
Open-closed groups	799.73	1	——
High-low ratings	4948.97	1	11.98 ($p < .01$)
Interaction	180.78	1	——
Between	413.23	16	

Ignoring the open-closed distinction, the association between ratings of the newcomer on "quality of ideas" and on "pleasantness" and amount of influence was also estimated by correlations techniques.

The relationship between attitude toward the newcomer and change of opinion in the direction of the newcomer's position was again most pronounced with reference to the Negro groups. The differences among these

TABLE 4
THE RELATIONSHIP BETWEEN ESTIMATE CHANGE AND THE RATINGS OF THE NEWCOMER CONTROLLING FOR THE RACE OF THE NEWCOMER

	Race of Newcomer	
Rating	White	Negro
a. Quality of ideas	0.43 ($p < .05$)	0.75 ($p < .01$)
b. Pleasantness	—0.08 (n.s.)	0.38 (p about .05)

correlation coefficients barely approach statistical significance ($.10 > p > .05$).

Finally, it was observed that the ratings of the Negro newcomer on both rating items tended to be *higher* (more positive) than the ratings of the White newcomer. A *t*-test of the difference between the means of the two distributions was significant at the 0.05 level of confidence.

D. Discussion

Clearly statistical support for the initial hypotheses is lacking. However, there is some evidence (see Table 1) that under the closed-Negro conditions the newcomer was not integrated with the host group members to the same degree as under the remaining experimental conditions. To this extent, the results are consistent with the original framework concerning the greater inter-personal orientation in closed groups and the relative lack of acceptance of a newcomer who is readily distinguishable as a representative of a minority ethnic group.

Yet, prejudice is necessarily parochial in nature. Thus, the results may simply reflect the relative strength of minority group effects in the given experimental environment. It is to be hoped that the experiment can be conducted in other settings since the techniques involved, particularly the objective measures of influence, may be readily adapted for cross cultural studies concerning racial or ethnic prejudice.

It seems possible that the results were attenuated by the nature of the experimental design which involved groups rather than individuals. Where it was possible to analyze the results in terms of individual behavior, the findings were psychologically and statistically significant.

The ratings by the naive members concerning the "quality of ideas" and "pleasantness" of the Negro or White newcomers were the key to these latter analyses. The newcomers, of course, were confederates whose messages to the naive subjects were essentially identical. Nevertheless, it was found that the relationships between the two critical ratings were significantly close only under the conditions in which the newcomer was a Negro. In addition, an analysis of variance of individual changes of estimate and attitude toward the newcomer also indicated a significant relationship, but again only with reference to the Negro newcomer. Moreover, correlations between ratings and change of estimate were substantial only with reference to the Negro newcomer. Finally, and quite unexpectedly, the results strongly suggest that the Negro newcomer was rated significantly higher with regard to "quality of ideas" and "pleasantness" than the White newcomer.

It would be difficult to ignore these findings and conclude that the naive subjects were unaware of the newcomer's race or to conclude that naive subjects reacted similarly with the Negro and White newcomers. Rather, the results seem to suggest that the panels which obliterated the physiognomy of the participants notwithstanding, some members were acutely aware of the physiognomic differences as cultural cues and reacted according to their ethnic biases.

It would appear that the Negro newcomer was categorized (1) according to a single salient dimension (race) and all other ratings were perceived as congruent with this dominant dimension. Perhaps, even the acceptance of the Negro newcomer's ideas depended to a greater degree on his acceptance as a "pleasant" group member.

Thus, the interpersonal perception of "pleasantness" would seem to be a more critical dimension with reference to individuals who are easily identified by some external differentiating characteristic. Moreover, this "halo effect" tended to pervade the entire group under the Negro conditions. Even the ratings of a White confederate who had entered the group on a task prior to the induction of the Negro newcomer tended to be significantly related ($r = 0.49$, $p < .05$) whereas this relationship faded with reference to the totally White groups ($r = 0.34$, n.s.). Thus, perhaps more generally, the results seem to suggest that interpersonal orientation is a more critical consideration in inter-racial groups or in groups where there are distinct differentiating characteristics among the various subgroups.

The unexpectedly higher ratings of the Negro newcomer in comparison with the White newcomer merit a brief comment even though the results are merely suggestive. Most citizens of the United States are aware of the race prejudice concept. It is probably also safe to assume that in college populations the concept has even become obtrusive in the presence of a Negro. Thus, it might be anticipated that a White college student's interaction with a Negro student tends to be less spontaneous, and perhaps over-controlled. Moreover, when faced with the necessity of rating a Negro on a given dimension, this interpersonal anxiety might be manifested by an attempt to "bend over backwards" to make an accurate appraisal. In so doing the Negro is actually rated higher than his White counterpart. In measurement terminology, the phenomenon may be described as a "generosity error" toward minorities or simply an overcorrection for possible prejudice.

E. SUMMARY

This laboratory experiment was concerned with the assimilation of the newcomer under conditions where the host group expected membership

changes (an open group) or did not expect membership changes (a closed group) and the perceived relative status of the donor group or culture (represented by a Negro or White newcomer). While the results with reference to the initial hypotheses were not statistically significant there was some evidence that under the closed-Negro conditions the newcomer was the least well assimilated. However, it was noted that the Negro newcomer and confederate tended to be rated significantly higher than the White counterparts. Moreover, the "halo effect" between ratings tended to be more pronounced with reference to the Negro newcomer. Finally, the relationship between attitude toward the confederate and actual change of opinion on a given task was more pronounced under the Negro conditions. It was concluded that interpersonal orientation is a more critical consideration in inter-racial groups or in groups where there are distinct differentiating characteristics among the various subgroups. The higher ratings of the Negro confederates were attributed to "overcorrection for possible prejudice."

REFERENCES

1. ALLPORT, G. W. The Nature of Prejudice. Cambridge, Massachusetts: Addison-Wesley, 1954.
2. BERTALANFFY, L. V. The theory of open systems in physics and biology. *Science,* 1950, **111**, 23-29.
3. CARTWRIGHT, D., & ZANDER, A. *Group Dynamics.* Evanston, Ill.: Row, Peterson, 1953.
4. CATTELL, R. W. New concepts for measuring leadership in terms of group syntality. *Hum. Rel.,* 1951, **4**, 161-184.
5. MURRAY, H. A. Explorations in Personality. New York: Oxford Univ. Press, 1938.
6. SHUVAL, J. T. Patterns of intergroup tensions and affinity. *Internat. Soc. Sci. Bull.,* 1956, **8**.
7. ZILLER, R. C., BEHRINGER, R. D., & JANSEN, M. J. The newcomer in open and closed groups (Unpublished report). Fels Group Dynamics Center, Univ. Delaware, Newark, Delaware, 1959.
8. ZILLER, R. C., & BEHRINGER, R. D. Assimilation of the knowledgeable newcomer under conditions of group success and failure. *J. Abn. & Soc. Psychol.,* (in press).

Fels Group Dynamics Center
University of Delaware
Newark, Delaware

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 85-104.

LSD AS AN ADJUNCT TO PSYCHOTHERAPY WITH ALCOHOLICS*

South Oaks Psychiatric Hospital, Amityville, New York

A. Rolo, L. W. Krinsky, and L. Goldfarb[1]

A. Introduction and Problem

The treatment of alcoholics in a hospital setting, particularly in regard to psychotherapy, is and has been one of repeated failure. The alcoholic seems to view the hospital structure as a punitive one, and most individuals suffering from this type of disorder seem to seek ways to avoid interpersonal involvement with members of the staff. For the most part their general approach to their problems seems to be, "I'll never do it again." There is, and has been, a continual difficulty in keeping alcoholics in a hospital for any appreciable length of time. It has been found that almost as soon as the acute signs of alcoholism subside; that is, as soon as the coarse tremors of the extremities and nausea are relieved, the alcoholic makes demands relevant to leaving the hospital. Often, from this point on, they are completely inaccessible to psychotherapeutic intervention. Outpatients are equally inaccessible. Unable to tolerate more than minimal anxiety, they often relapse between therapy sessions.

A chemo-psychotherapeutic technique which would be helpful in making individuals more accessible to dynamic interpretations would have to satisfy four criteria: (*a*) it must have properties which would lessen the inhibition of the individual; (*b*) it would have to permit the therapist to be able to move through the defenses and allow a rapid penetration to the unconscious; (*c*) it would have to have a lack of toxicity, and permit the patient to return to normal functioning; (*d*) it should have properties which would not make the patient amnesic for what had transpired during the conference.

LSD, D-Lysergic Acid Diethylamide, has been the subject for experimentation since 1943. Most of the studies reported, however, have concerned themselves with general pharmacology. Very recently the use of LSD as a psychotherapeutic adjunct has been described.

In the recent past, Busch and Johnson (4) have concluded that LSD may

---

* Received in the Editorial Office on April 23, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Physician in Charge; Chief Psychologist; Clinical Director, respectively. LSD-25 supplied by Sandoz Pharmaceuticals.

offer a means for more readily gaining access to the chronically withdrawn patient. They have further stated that LSD may serve as a new tool for shortening psychotherapy. Frederking (5) discussed the therapeutic effects of LSD and describes the therapeutic effect as being such as to aid psychoanalysis in 24 cases. Abramson (1), Frederking (6), and Langner and Kemp (7) described experiments in which they conclude that LSD both enhances and facilitates psychotherapy by allowing a rapid penetration to the unconscious. Abramson reported that under the effects of LSD the therapist can recognize the presence of the processes of ego reinforcement, and that this medication operates effectively to allow the patient to utilize preconscious and unconscious material during therapeutic interviews lasting as long as four hours, and covering periods of observation lasting six hours.

Benedetti (3) reported in regard to the treatment of alcoholics with LSD. He discussed a marked psychocathartic effect resulting from the taking of this medication. Arnold and Hoff (2) have also worked with alcoholics, but do not report making therapeutic progress. Nunes (8) pointed out that LSD has non-harmful and non-toxic effects with alcoholics. Smith (9), in employing LSD with 24 alcoholics, concluded that the medication in conjunction with psychotherapy gives favorable results with a group of character disorders and some psychopaths, but that borderline or actual psychotics did not appear to benefit. He suggested that more extensive trials be made with this medication. On the basis of the literature reported, it seemed likely that LSD might offer a unique approach in facilitating psychotherapy with alcoholics.

B. PROCEDURE

The patients in this study were 12 alcoholics, eight female, four male, who have a history of chronic alcoholism ranging from 8 to 20-plus years. Alcoholics were defined as: (*a*) patients whose drinking, either periodically or daily, is such as to interfere with their productivity; (*b*) patients who, during their periods of drinking, begin in the early morning, continue throughout the day, and imbibe at least one fifth of whisky daily; (*c*) patients whose drinking is accompanied by a marked degree of avitaminosis; (*d*) patients who may or may not develop psychotic symptomatology—i.e., DT's; (*e*) patients who have some degree of liver damage related to their excessive drinking.

LSD-25, D-Lysergic Acid Diethylamide (Sandoz), was used in all cases. Each patient had been amply prepared for this interview, was given notice as to when it would be conducted, and it was undertaken only after each patient had given his or her consent and had expressed a desire for aid. Ten patients were started on 100 mcg. of LSD which was dissolved in a half cup of water or orange juice. After one and a half to two hours had elapsed, two of these patients were given an additional 100 mcg.; one was given an additional 50 mcg.; two patients received 50 mcg., and one of these was administered an additional 50 mcg.

This was administered on the basis of experience, inasmuch as it was found that this increased as well as hastened the onset of the desired reaction. The reaction which was desired was to make the patient more accessible, but was not of such intensity as to create hallucinations.

Each patient was seen in a bedroom-sitting room in the cottage in which he or she resided. The members of the professional staff who were present in the extended conference were known to the patient, and at the beginning of the interview, as well as at any time during it, the patient had the option of excluding anyone that they so desired.

The method that was found most productive was one in which at least two, and preferably three, members of the professional staff, as well as a registered female nurse, were in attendance. It was found that having at least two doctors present allowed for a much more intensive approach than would have been possible with only one member of the staff. The patients not only did not seem to mind the multi-therapist approach, but often expressed their gratefulness to members of the staff for their interest. At the conclusion of the interview, they frequently expressed their belief that it had been so successful because of the participation of so many staff members.

In this novel approach, one member of the staff would interview for a certain period. The others would make notes of material which was introduced by the patient, and which might prove fruitful if explored. After some time had elapsed, the other members of the staff would gradually and almost unobstrusively take over the questioning. It was found that this interview technique, which ran from a minimum of three to a maximum of six and a half hours, was so great a strain, that each member of the staff who attempted it alone found his concentration powers were not as keen after a few hours. It was also found that having trained observers, both with psychological and psychiatric backgrounds and hence with different orientations, allowed for moving into areas which one member of the staff would not have gone into alone.

The R.N., who sat unobstrusively to the side, took notes relevant to the time the interview commenced, the physical changes reported by the patient, and the time of the conclusion of the conference. Many patients later reported that they felt more secure with a female in the room (possibly a mother substitute), and at the conclusion of the interview the nurse remained with the patient for at least one to two hours. Having someone with him who had been involved in the interview served as a distinct source of comfort to the patient. This opinion was gained from subjective reports by the patient.

Another aspect of this technique which was found most helpful involved having lunch with the patient. Since the conferences usually began at 9:00 to 9:30 A.M., having luncheon served in the bedroom-sitting room with the patient seemed a natural consequence. This served greatly to enhance rapport, and during luncheon the discussion moved away from areas of conflict. A few patients refused to rest, and felt the need to continue to ventilate. Immediately afterwards, dynamic material was discussed again.

Utilizing the method which was developed, the discussion was opened prior to giving the medication, re-emphasizing to the patient that there would be some pleasant physical sensations, and that it would make the patient feel more comfortable in speaking. After the medication was taken, and before it began to take effect, a general discussion was engaged in relevant to the amount of intake of alcohol.

Within 15 to 30 minutes after taking the LSD, the patient began to describe physical sensations (tingling of extremities and light-headedness), and the interview moved into having the patient discuss his or her early childhood.

By the end of one to one and a half hours, the patient usually reported a feeling of extreme well being, a feeling similar to that they said they experienced from one or two cocktails. At this point they were usually mildly euphoric and often seemed pressured to speak. A few patients reported some slight disturbance in vision, a form of diplopia, but no hallucinations. At this point the patient was led into more dynamic material relevant to parent/child relationships.

In the last hour to hour-and-a-half, the dynamic material was discussed again, and by the last hour of the interview the patients were able to make their own interpretations. Sudden insight reactions ("Ah Ha") were frequently encountered. During this time interpretations would be made by one member of the staff, enhanced and clarified by another staff member, and often the patient would then take over and interpret to himself or herself. In the last half hour, the material was reinterpreted, and the patients were allowed ample opportunity to discuss their feelings, newly formed attitudes, etc.

Termination of the interview was mutually agreed upon when it became evident that further material would not be forthcoming. In all but one instance, the patients were most grateful, expressed the belief that they were "free" or "cleansed." Such attitudes carried over on the following days, when the patients were eager to meet with staff members to have material reinterpreted. The interpretations which were arrived at almost invariably revealed the basic problems of each patient which were directly related to the excessive use of alcohol.

C. RESULTS

In all cases, the technique employed resulted in the patient bringing forth much of the dynamic material relating to the onset of his or her alcoholism. In all but one case, the patient was able to utilize this material in an insightful condition. They were able to relate their alcoholism to their own problems and adjustment, and in the majority of the cases spontaneously verbalized exception, the dynamic material was brought forth, and in spite of the fact that it was interpreted and reinterpreted, the patient remained highly resistent. He said that he could see no relationship whatsoever between his drinking and the dynamic interpretations. He "promised" that he would be able to stop drinking, "only by the force of my personality."

The technique employed is not viewed as a curative measure, but only as a technique which permits rapid psychotherapeutic intervention with alcoholics. In view of the small sampling and the short period of time in which this technique has been employed, these findings should be viewed as tentative. It appears, however, that this technique has been successful in 10 of the 12 cases in which interviews were conducted. In regard to the other two cases, one was found to be suffering from well-integrated paranoid delusions, and was subsequently diagnosed as schizophrenia, paranoid type. The other failure has not as yet relapsed, has been home 10 weeks without drinking, but during the conference revealed himself as suffering from a severe compulsive personality disorder. He announced at the end of the interview that he had learned nothing about himself (although a great deal of dynamic material was unearthed), and angrily left the hospital two days later, reiterating his belief that the hospitalization had been detrimental to his health and that he would stop drinking, "by the force of my personality."

1. *Case 1*

A 49-year-old female clerical worker who has a history of alcoholism dating back 15 years, being described as a "silent wine drinker." She gradu-

TABLE 1

A SUMMARY OF THE RESULTS ON THE PATIENT POPULATION

Patient	Age (in yrs.) at time of interview	Number of previous hospitalizations for alcoholism	Length of time patient has been recognized as alcoholic (in yrs.)	History of psychosis due to alcoholism	Dosage of LSD (mcg)	Results (in terms of depth of understanding achieved by patient)
1	33	4	8-12	Yes	100	Excellent
2	50	1	9	Yes	50/50	Excellent
3	33	0	5	No	100	Fair
4	42	0	10-15	No	100	Excellent
5	56	2	15-20+	Yes	100	Poor
6	56	3	5-10	No	100	Excellent
7	57	9	20+	No	100/100	Good
8	50	4	15	No	100/50	Excellent
9	53	2	10	No	100/100	Excellent
10	49	0	15	No	100	Excellent
11	35	6	17-20	No	100	Excellent
12	53	2	8	Yes	50	Poor

ally progressed to the point where she was drinking a quart of whisky daily. Whenever she was able to get it she would become extremely overtalkative, nasty, and sarcastic. She would walk about the house muttering to herself, fall asleep for a short time, waken, and begin drinking again.

Her birth and early childhood were described as "normal." She has been in generally good physical health with the exception of having an appendectomy as a child, pneumonia in 1956, and liver damage as the result of her drinking. Her relationship with her parents and siblings was said to be a wholesome and well-adjusted one. She was a high school graduate, attended college, but did not graduate. She had worked in the same office for 21 years, and had obtained a leave of absence in order to enter the hospital. She was described as never having had much social life, and no hobbies. She was married for a few years; however, her husband died of cardiac disease. She had one son, 15 years of age, who lived with her in her parents' home.

An interview under LSD was conducted with Mrs. F. on November 25, 1959. Those present during the interview were three members of the staff and an R.N. The interview started at 9:30 A.M., and within 20 to 25 minutes the patient began to have a reaction and was able to verbalize freely.

A great deal of dynamic material was uncovered. The patient came to understand that she has been totally incapable of expressing hostility to anyone. It also became evident that she has internalized much of her feelings of hostility, defending against their outward expression by a retreat to religion.

Her hostility seems to be directed mainly against a younger sister and her mother. The patient's mother apparently delegated the care of all of the siblings to Mrs. F., she being the oldest. The patient consistently described her mother as "being around" and implied as well as stated frankly that during times of crisis or periods when her mother should have been about to be responsible for the care of the children, she had gone to the movies, shopping, was "in the kitchen," and left all responsibility to the patient.

It appears that this taking over the mother role was reinforced by her father, for whom she expresses great love. She feels that she was his favorite child, and he is certainly her favorite parent. But he, too, would confer with the patient in regard to the care and management of her siblings. During this conference she came to understand that he had cast her in the role of mother and that over the years she has lived this role.

This extended to the point where all the significant relationships she has had with strangers have been those in which she can be motherly. Prior to meeting her husband, she had had a deep friendship with a cerebral palsied young man she had to care for. She married her husband, a victim of encephalitis, knowing that he was ill with hypertension, and apparently was able to exercise a maternal role in regard to him.

Her hostility towards her sister is of long-standing duration. The patient's sister is evidently more beautiful than she, and although only 15 months younger, she was spared having to care for the younger siblings as did Mrs. F. Several years ago her sister was employed in the theatrical business, apparently kept late hours, drank to excess, was sexually promiscuous, and subsequently became pregnant. The child was delivered in the patient's mother's home. Instead of her mother taking care of the necessary details, Mrs. F. was summoned from work, and even left her husband and child for a few weeks while she cared for her sister's child. She abreacted as she said she helped the physicians when they were transfusing her sister, helped make arrangements in an effort to deceive the neighbors that her sister had been married, took the steps necessary to see that the child was legally adopted by her mother and father, and also went so far as to be involved in seeing that the father of the child made a $10,000 settlement. Afterwards, her sister took her to live with her in the home she was able to buy with the $10,000, repeatedly and in a belligerent manner would inform the patient that she was living in her home, and eventually "threw" her out. During this time her sister bought a car, could not drive, and forced the patient into the role of chauffeur.

In spite of this, the patient claimed during part of the conference that she loved her sister, was grateful to her, and bore her no hostile feelings. Eventually, as we were able to deal on a deeper level of the unconscious, the patient verbalized extreme hostility towards her sister, and although wincing as she did so, called her a "tramp" several times.

A relationship between the patient's drinking and her underlying dynamics was uncovered. She came to see that drinking has a threefold purpose for her; that it represents an internalization of feelings of hostility, that she figuratively drinks her anger, and in this way convinces herself that she is worthless and a "bum." Drinking, then, is a self-derogating device in which the patient punishes herself for feelings of hostility which are harbored on an unconscious level. A second reason for drinking is related to her need to take a dependent role. The patient, who recognizes that she never had an opportunity for a normal childhood, has, by virtue of her drinking made it necessary for her mother to take care of her and at the same time, she, the patient, is able to divest herself of responsibilities and of taking care of others as she always did. During her long history of alcoholism, her mother has had to care for the patient's child while the patient was "around." Mrs.

F. came to see that this drinking also served as a device for her to express hostility covertly.

By the conclusion of the conference, Mrs. F. had been able, in an insightful manner, to make these associations and accept the interpretations. She was greatly relieved, seemed genuinely moved by much of the material that was unearthed, and said in a very sincere manner that she did not think she would have to drink again. She added that the revelations were too new to her to be able to make any forthright decision. She was told that she would be seen by a member of the professional staff who had been at the conference, to go over these interpretations with her, and she was also urged to seek psychotherapeutic aid on the outside. She came to understand that this need to be maternal has permeated all life relationships and thereby impaired her functioning and life adjustment.

2. *Case 2*

A 35-year-old male insurance broker, who has been married for 7 years and has three children, 5, 4, and 2 years of age. This patient is reported to have been drinking for the last 20 years, and heavily off and on for the last 17 years. For the 5 - to 10-year period prior to his admission to the hospital, he was said to be a "continuous drunk," and would drink more than a quart of whisky daily. In spite of his joining Alcoholics Anonymous, as well as taking antabuse, he was unable to remain sober for more than a five-month period.

He had several previous periods of hospitalization in general hospitals, rest homes, and homes set up solely for the care of alcoholics. In the two-year period prior to his admission to this hospital, his drinking became progressively worse, and he would drink every day. On occasion he would take up to two quarts daily.

The patient's familial relationships, particularly with his father, have been difficult. The father was described as very hard to get along with, and apparently there were many arguments in the home between the mother and father. His mother is described as overly protective, very close to the patient, and a woman who has always projected the cause of her son's alcoholism onto others. This patient is one of three siblings. His older sister is a heavy drinker, who takes a great deal of wine each night, stating that she must calm her nerves after the children have gone to bed. His younger sister is married to a heavy drinker, but apparently does not drink herself.

He is a graduate of a military academy, during which time he was very active in sports. He was given a scholarship to a Jesuit college, and graduated with a degree in economics and philosophy. He worked for his father in his insurance business for approximately seven years, but was never happy there because of a personality conflict. Two years ago he decided to open his own business, did very well in the beginning, but as his drinking has become more of a problem, all the responsibility of the management of the office has been taken from him and his wife takes care of all business matters. This patient is said to be somewhat withdrawn, with no close friends, likes to garden, and enjoys doing charcoal paintings and hunting.

An LSD interview was conducted with Mr. O. on December 2, 1959. Those present at the interview were three staff members and an R.N. The interview was started at 9:30 A.M. and terminated at 1.25 P.M. The patient was given 100 mcg. of LSD at the time the interview began.

In accordance with the plan which was evolved with the use of this medication, the patient was encouraged to discuss his history of alcoholism in the period immediately after he took the medication and before it had begun to take effect. He discussed the fact that he began to drink at the age of 15, that he took whisky at that time, was castigated by his father, who told him he was too young to drink whisky and should drink beer. He related that the drinking continued throughout the time he was in the Army, when he was in college, and during his marriage. He said that he had been drinking to excess within the last several years, and has learned from a bartender that he would drink 40 to 50 glasses of beer a day. He also admitted to drinking at least a fifth of whisky daily.

As he began to experience some reaction to the medication, the second phase of the interview was entered. During this time the patient discussed his early childhood, parental relationships, and attitudes towards his siblings. He related that his home life had been a stormy one, that his mother was "vicious" in her continual domineering attitude towards his father. He related that she picked on his father almost mercilessly and that she used him (the patient) as a pawn in the battle with her husband. The patient feels that his mother overindulged him and always met all of his needs prior to taking care of his father, but that she did so not because of genuine love but because this was a way for her to express hostile feelings. He related that when his father was not about, his mother would be vicious towards him, that she would laugh at him in a most derogatory manner and made him feel inconsequential and rejected.

During this second phase of the conference, he discussed his father in an ambivalent, almost confused manner, spoke of hating him at first, and with tears in his eyes amended this to say that he did not really feel this was genuine hostility. He spoke of his siblings in a rather general fashion, indicated that they were more subject to hostile expressions from his mother, but it did not seem that he was genuinely involved with his sisters.

In the third phase of the conference, material relating to parent/child relationships was entered into on a deeper, more dynamic level. The patient came to understand that he has unconsciously been trying to identify and compete with a father whom he has always felt was more successful, more intellectually endowed, and more social than he. His father has always discouraged the patient about accepting the idea that he is an alcoholic. When he was 16, he advised him that he should learn to drink beer before taking whisky, and regaled him with stories of his own drinking during his teenage years. Of late, he has again told the patient that he is not an alcoholic, and should not be in A.A., but that he should "learn how to drink."

The patient's idea regarding the prestige of his father is related to observing his father drinking with his many friends who were also alcoholics. He has been able to rationalize his father's weaknesses in dealing with the mother by coming to view them as a virtue. He relates that his father has always said, "When they bury me, they'll put on my tombstone, 'Here lies a man who had infinite patience'," and he feels that this is an indication of his father's strength rather than weakness.

In this identification and competition the patient has always felt, and apparently always *has* come out second best. For example, his father is a graduate of an "Ivy League" college, and made it quite clear to the patient on his graduation from college that he demeaned the patient's accomplishment (our patient graduated from a small Catholic university). His father is in the insurance business, and the patient entered this business, but again with less success. As a matter of fact, his drinking has made it impossible for him to carry on the business and he has turned it over to his wife.

In his marital relationship, although his father was clearly subservient to his mother, the patient felt that his father was still able to maintain an outward appearance of being the "man of the house." He recognizes his own deficiences in this regard, sees that his wife has taken over the business and control of the finances, that she gives him an inadequate allowance when he works and nothing at all when he drinks.

In his social relationships the patient has again tried to emulate his father, but with less success. The only friends he has had are those that he meets in bars, who are also alcoholics, and in order to impress them with his importance he begins to talk about his accomplishments as a sports star and as a soldier in Japan during World War II. (His father was evidently able to attract friends without having to resort to resumés of his past accomplishments.)

His competition with his father and desire at the same time to gain his favor extends to his career in sports. The patient was active in baseball and lacrosse, but was very desirous of playing basketball. He was encouraged in this by the coach of the team, however, he relates that his father forbade him to do so, stating that if one played basketball in a closed arena one was subject to tuberculosis.

During this interview the patient came to realize that courting the favor of his father as he did was the very reason that he did not play basketball, that he acted as a young child even though at the time he entered college he had completed Army service and was in his 20's. During the early part of the conference the patient had related that his nickname was "Joe." When questioned about this during the third phase of our interview, he revealed that he was overjoyed to take the nickname Joe because he could then give up his own name, Charles, Jr. He realized at this point that the idea of being "Junior" to his father was abhorrent to him. Suddenly and in a highly insightful manner, the patient, came to realize that his life pattern has been one in which he has either served as a buffer between his parents, or has unconsciously tried to emulate his father. He said, "I have never been me . . . I have got to be me." From this point on the material which was revealed was reinterpreted to him.

At the conclusion of the conference he seemed to have good understanding of what he learned and spoke of his hope and resolve that he would never drink again. He was encouraged, and seemed to understand the necessity for his continuing as an active member of A.A. At the same time he came to understand that he was in need of intensive psychotherapy for a protracted period so that the understanding that he developed during this conference could be enhanced and so that he could see how it had affected and permeated all of his life relationships. The patient seemed exceedingly grateful to the members of the staff who had been in the conference with him, expressing his thanks for the aid that he had received and at this point the interview was concluded.

3. *Case 3*

A 32-year-old married female, who has three children. This was her third period of hospitalization in a mental institution within a one-year period and on her arrival she was reported to have been drinking to excess for six or seven years. During this period she had had six or seven admissions to the psychiatric division of a general hospital, had also been in hospitals set up solely for the care of alcoholics, as well as numerous "homes" for alcoholics. This is a patient who was said to consume one and a half to two quarts of whiskey daily, and who has been unable to stop in spite of an association with A.A. During periods when she drinks, she begins in the early morning, filling a coffee cup three-quarters with whiskey, one-quarter with coffee, and then continues throughout the day. She does not eat during these periods, is greatly undernourished, and is reported to have liver damage. This is a patient who hides liquor in the home, orders it through her grocer, butcher, cleaner, etc., and during a period of one month has charged as much as $140 for whisky. During times in the past, as well as during her first hospitalization here, psychotic symptomatology was noted.

Familial relationships were described as healthy, the patient's father died when she was only 4 or 5, was said to own a large contracting company, was an honorary member of the police force and very well respected in the community. Her mother, who is 61 and resides with the patient, was described as a pleasant, easy-going individual, who has always been very close to the patient. Her husband holds a very responsible job with a manufacturing company, and is reported as intelligent, pleasant, with a good sense of humor, a man who is devoted to his wife and children.

This patient attended parochial schools in Chicago, Ill., and graduated from high school with barely passing grades. She had a great many friends during the time she went to school, has since given them up, and most of her associations are with people who have the same problem as she. She has no hobbies, "only one, and that is her drinking."

An LSD interview was conducted with this patient on November 13, 1959. The interview was begun at 9:15 A.M. and was concluded shortly after 1 P.M. Those present during the conference were two staff members and an R.N. The patient was given 100 mcg. of LSD at 9:15 A.M. Prior to the time she took the medication, the patient, who seemed apprehensive and tense, asked that we address her by her given name, Mary. She expressly asked that we not refer to her as Mrs. J.

At this point, utilizing the technique that has been developed, and prior to the time the medication began to take effect, the patient was encouraged to discuss her drinking. She did so rapidly, pointing out that she began to drink even prior to the time she was a teenager. She discussed the fact that her mother was always having parties and that she and her brother would entertain themselves after the parties, drinking the remainders of Manhattans and taking the cherries in each drink. She says that by the time she was 13 or

14 her mother would have a scotch and soda waiting for her when she came home from school, and that the two of them would drink together, to help pass the time. The patient said that she now understands why she had so much difficulty with her school work, inasmuch as she could not concentrate after having had two or three drinks. She related, too, that her entire family has always drunk freely and openly, she has always been surrounded by drinking people and a great deal of liquor. She says that she began to drink heavily when she was 18 or 19 and that this has continued over the years. Within the last several years, alcoholism has become a problem and she can take even more than a quart of whisky daily.

As she reached this point, she began to experience a reaction to the medication, spoke of a tingling sensation in her extremities and of a feeling of light-headedness. At this point she was diverted into a discussion of parental relationships. The patient hastily dismissed the topic of her father, pointed out that he had been a "gangster" who had been murdered by an opposing gangster faction, that this was a topic that had always been taboo within her family. She said she did not remember how her father looked or acted, and it was quite evident that it was anxiety-provoking for her to continue.

Her discussion of her mother was a more elaborate one, and one which was characterized by her expressing marked feelings of hostility. She related that her memory of her mother was that of a "queen," that she recalls her mother remaining in bed late each day, and that both she and her brother were only permitted short visits with her in the afternoon. The patient remembers with some feelings of disgust that on weekends she was permitted to go to her mother's room for a longer period and to share sections of the newspaper with her mother while they both read.

This image of her mother as a "queen" extends to the point where she saw her mother inviting a great many people to the home for drinking parties and avoiding all interpersonal relationships with the children. She relates that instead of being given love and maternal affection, she was given unlimited access to charge accounts to buy whatever she wanted, keep whatever hours she wished, going with whomsoever she desired, and to bring anyone that she wished home.

By this point in the conference it was possible to move into deeper dynamic material. In questioning her it became evident that the anxiety regarding her father is the result of most complicated feelings. Abreacting, the patient was able to recall a most pleasant relationship with her father, who must have loved her deeply and whom she loved in return. Although he may have led a most unsavory life, the patient's memory of him is of a kind, sweet, nondrinking individual who would lead a horse about on the grounds while she was sitting in the cart. She was also able to recall the details of his death, of how he left the home one morning, there were shots, the maid ran to the door, ran back in, her uniform covered with blood and said, "Children, pray for your father, he's dead." The patient was not permitted to go to the funeral and apparently in order to escape the same gangsters who killed her father, she, her mother and brother had to move from one relative's home to the other for several months.

Her memory of her mother as a "queen" dates from this point, and it was suggested to the patient possibly her mother became deeply depressed at this time and may have been ill, this resulting in her withdrawal rather than in conscious wishes on her part. The patient came to feel that she has always held her mother responsible for the death of her father. She thinks that her mother was actively involved in the sale of alcohol during prohibition, with her father. She has been led to feel over the years that the reason her father acted illegally was because his wife had made so many financial demands upon him. The problem that the patient has had in regard to her father is that she has always loved him, but at the same time has been made to feel guilty that he was her father. In school it was always brought up to her that she was the daughter of a gangster, children shunned her, and she feels the nuns who taught her were very critical and cruel.

At this point it was possible to introduce into the interview her statement about wanting to be called only by her given name. This seemed to introduce a vast area of information. The patient said that she has always quickly made friends with others, so that she might call them by their given names, and that having a surname is most anxiety-provoking. She relates that she never wanted to be known by the name of Mary R., this being her maiden name, as her father was infamous and it brought on the anxiety about having a gangster father. This has extended to the point where she does not use the name R. as a middle name, but instead uses the name Sarah.

Her unwillingness to use the name J. is the result of guilt she has in this regard. She came to understand that she married at the young age she did, to a man she did not love, because he offered her an escape from the name R. Using her present surname is most anxiety and guilt-provoking, and she prefers not to do so.

The patient came to see that her drinking represents a form of hostile expression for her which can be used against her mother. It seems that on a deep level of the unconscious she has introjected the idea that she is a "gangster's child" and drinking allows her to act in this fashion. The principal reason for the drinking, however, is clearly that it allows for the expression of hostility when intoxicated. Even here she sees that she has been unable to bring up her true hostile feelings towards her mother, always passing out before she is able to discuss her late father.

As the patient came to understand this, as the material was reinterpreted to her and as she herself experienced the "Ah-Ha" reaction, she came to see that there was no need for her to drink in the future. It was impressed upon her that she requires not only active participation in A.A. after she left the hospital, but should also have the benefit of extensive psychotherapy to further explain these interpretations and to see how they have generalized and permeated all of her interpersonal relationships. The patient seems most willing to take such aid in the future, and in a rather sardonic manner expressed her belief that she now would no longer drink because there would be no reason for her to do so.

4. *Case 4*

A 56-year-old female patient with a history of heavy drinking over the last 10 years, and to great excess within the last five to six. She was previously hospitalized at Knickerbocker Hospital and later at another hospital under the auspices of A.A., where she remained for two weeks. The patient is reported to drink at least a quart of whisky daily, is said to become most unattractive when she drinks, neglects her home and person, becomes argumentative, uses profane language, and the problem has become so acute that her husband has now separated from her and will not return to live in the family home. This is a patient who begins drinking in the early morning, usually has about 10 cans of beer and then starts to drink whisky by midmorning. She is reported to have some liver damage as a result of this drinking. This is a patient whose birth and early development were described as normal, whose relationships with her parents were said to be happy, the family being a very close-knit unit. She is the youngest of seven siblings, one of whom has been hospitalized for 30 years in a State hospital, diagnosed dementia praecox, and the remainder are said to be well-adjusted emotionally.

The patient is married for 33 years, her husband is said to be an individual who has a great deal of finesse, but who can also become brutal and assaultive. He is a heavy drinker, encourages the patient to drink, wants her to entertain his business contacts, but has no patience with the fact that she cannot control her drinking. She is the mother of four children, all of whom are said to be in good physical health and well-adjusted emotionally. This is a patient who is a high school graduate, and also took a B.A. degree in education at one of the City colleges. She has been teaching off and on throughout the years of her marriage, and has apparently done extremely good work. She has had a very active social life, but mostly through her husband's business associations, and she has practically no close friends.

An LSD interview was conducted with Mrs. Y., on November 2, 1959. Those present at the interview were three staff members and an R.N. The interview was commenced at 9:55 A.M. and concluded at approximately 2:20 P.M.

Just prior to the time she took the medication, as well as for approximately the first half hour to 45 minutes afterwards, the patient discussed her history of alcoholism in a relatively superficial manner. She had not yet had any effect from the LSD and spoke of the fact that she had always been a social drinker, but that in the last several years she had been unable to control this. She said she drank more and more, to the point where her husband became disgusted with her and subsequently left.

She seemed somewhat hostile towards her husband during this phase of the conference and indicated that she felt that he was as responsible as she for her alcoholism. She said that he drank to excess himself but that he could manage his intake and she related that he encouraged her to drink. She claimed that her husband's business position was such that he must do a great deal of entertaining, that he does not want her to stop drinking completely, but merely to be able to control what she does take. She said that there had been times in the past when she had tried to stop drinking, but that he would almost deliberately taunt her by drinking fairly large amounts in her presence, often encouraging her to do so.

At this point in the conference she entered the second phase when she began to experience a tingling sensation of her extremities and a general feeling of light-headedness and well being. At this time she was diverted into a discussion of her early childhood. At first she was inclined to gloss over anxiety-provoking material and confine herself to a discussion of her own accomplishments in life. She spoke glowingly of her attributes, mentioned that she had received a college education under trying financial circumstances, that she had become a highly sought after teacher and that in recent years she had served as a substitute at the high school in her town. She spoke of her accomplishments as a mother, mentioned that she had four children, all of whom had had the benefits of college training, and with whom she feels she is a good mother. She spoke of her own home life as a child as having been happy, mentioned that her father and mother were hard-working people who loved each other.

By this point in the conference she was under the effects of LSD to the degree where it was possible to broach deeper material. She had given hints throughout the discussion that her need to gloss over and present material in such a favorable light was related to the fact that she had to defend against underlying anxiety occasioned by unpleasant experiences.

The patient began to discuss her father as a tyrant, spoke of him as having been of the lower class, brutal, punitive and rigid, a "typical German." She mentioned that her father's needs were always met before anyone else's in the home, that if there was only one piece of meat the children would sit at the table and watch him eat it. She recalled that he was assaultive on occasion, striking the children and her mother, that he drank to excess in "saloons" but that if anyone tried to urge him to stop drinking he would become brutally angry. She described her mother as a "perfect soul" and as she went on it seemed that she thought of her mother as a self-sacrificing, angelic, hard-working, completely submissive female.

It was possible to bring her into this discussion on an even deeper level and what eventually evolved was that the patient was not panicked but rather fascinated and sexually attracted to this brutal father. She was able to recall both in pre-adolescence as well as adolescence that she would lie awake until late at night, awaiting her father's return from the saloon. At first she said that she wanted to hear if he would be assaultive to her mother, but later, and evidently under great strain, she confided that she would wait to hear if they had sexual relations. She said their bedroom was next to hers and vividly described the sounds that her parents made while engaging in sexual relations. She spoke of her mother crying out as if in pain, "Stop, you're hurting me," or, "It's too hard," and of her feeling that her father was not going to stop. Abreacting, she came to understand that this was not revulsion she was feeling, but tremendous attraction and she suddenly also recalled that she did not feel her mother was in pain but rather in ecstasy.

The patient's own drinking is clearly related to her sexual needs. She described her husband as a mixture of "earthiness" with "ethereal qualities." She spoke of the sexual adjustment as having been fine, but expressed fears that it was becoming less satisfying as the years went on. The patient came to understand that her drinking increased during this period because her husband's so-called earthy qualities were diminishing as he was moving up the executive scale.

At this point she recognized she began to drink and would then begin to frequent "saloons" rather than "cocktail lounges," that she picked up stocky, coarse, crude men (father images) and that on innumerable occasions she engaged in sexual relations with them. Drinking thus became the only acceptable way that she could satisfy her own sexual desires for a father image and still maintain her integrity. The drinking increased in recent years because of the patient's distorted belief that as her husband moved up the executive scale he was becoming less and less like her father and more and more effeminate. She was able to unearth the distortion she felt that people in the upper echelons of executive ranks were always effeminate and less sexually interested than working men—i.e., her father.

This material was interpreted to her on several occasions, the patient seemed to experience the "Ah-Ha" reaction, spoke of a feeling of having had a burden lifted from her, and expressed her belief not only that day but on days following that she no longer had any need to continue drinking. The need for further psychotherapy was impressed upon this patient at that time and it is believed that she came to understand her need for further aid. She promised she would avail herself of further psychotherapeutic services once she left the hospital.

D. SUMMARY

A group of 12 patients with a history of alcoholism dating from 8 to 20-plus years were treated with LSD-25 as an adjunct to psychotherapy. The medication was given as an aid in establishing rapport and to allow the working through of material on the deeper levels of the unconscious. It was not intended as a curative measure. The principal value of this technique is that it allows a rapid working through of basic conflicts. In the past, hospitalized patients proved highly resistive to therapeutic intervention and out-patients, unable to tolerate the anxiety these conflicts provoked, relapsed into alcoholism before they could make appreciable progress.

The results of a useful and novel interview technique are described. Of 12 subjects in which this technique was employed, the results were favorable in 10. One of the subjects was subsequently diagnosed as schizophrenia, paranoid type, and the other as a severe personality disorder.

This study is part of a preliminary trial. It is suggested that these results justify more intensive investigation.

REFERENCES

1. ABRAMSON, H. A. Lysergic acid diethylamide (LSD-25): III. As an adjunct to psychotherapy with elimination of fear of homosexuality. *J. of Psychol.*, 1955, **39**, 127-155.
2. ARNOLD, O. H., & HOFF, H. Untersuchungen uber die Wirkingsweise von Lyseräurediäthylamid: I. Mitteilung. (Investigations on the mode of action of LSD: First communication.) *Wien Ztschr. Nerven.*, 1953, **6**, 129-150.

3. BENEDETTI, G. Beispiel einer strukturanalytischen und pharmakodynamischen Untersuchung an einem Fall von Alkolholalluzinose, Charakterneurose und psycholeaktiver Halluzinose. (Example of the study of the mental structure and of the pharmacodynamic investigation of a case of alcoholic hallucinosis, character neurosis, and psycho-reactive hallucinosis.) *Ztschr. f. Psychother. u. med. Psychol.,* 1951 **1**, 177-192.

4. BUSCH, A. K., & JOHNSON, W. C. LSD-25 as an aid in psychotherapy. *Dis. Nerv. Sys.,* 1950, **11**, 241-243.

5. FREDERKING, W. Ueber die Verwendung won Rauschdrogen (meskalin und Lysergsäurediäethylamid) in der Psychotherapie. (The use of mescaline and LSD in psychotherapy.) *Psyche,* 1953-54, **7**, 342-364.

6. ———. Intoxicant drugs (mescaline and lysergic acid diethylamide) in psychotherapy. *J. Nerv. & Ment. Dis.,* 1955, **121**, 262-266.

7. LANGER, F. W., & KEMP, M. D. LSD-25. *Quart. Bull. Pinebluff Sanit.,* 1956, **1**.

8. NUNES, E. P. Investigacoes com a dietilamida do acido lisergioco. (Studies on LSD.) *J. Brasil. Psiquiatr.,* 1955, **4**, 407-418.

9. SMITH, C. M. A new adjunct to the treatment of alcoholism: The hallucinogenic drugs. *Quart. J. Stud. Alcohl.,* 1958, **19**, 406-417.

South Oaks Psychiatric Hospital
Amityville, New York

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 105-110.

A CASE OF EXPERIMENTAL NEUROSIS AND RECOVERY IN RELATION TO THE ORIENTING RESPONSE*

University of Arkansas Medical Center and Johns Hopkins Medical School

Roscoe A. Dykman and W. Horsley Gantt[1,2]

This report is concerned with the development of an experimental neurosis in one male dog (Schnapps) and his recovery as brought about by mild electric shocks. While this dog has been mentioned elsewhere (7), no detailed report of his behavior has been made. It is believed that the behavior is of sufficient theoretical value to warrant a complete report.

Schnapps was one of four dogs undergoing preliminary training for an autonomic conditioning experiment (1, 2). The preliminary (preconditioning) training consisted of (*a*) habituation to the experimental room and the Pavlovian type stand, and (*b*) orienting training.

Any mild or moderate stimulus causes an animal to direct his appropriate receptors into the most favorable position for receiving the stimulus. In the case of an auditory stimulus, the dog turns his head in the direction of the signal and pricks up his ears. There is nearly always a noticeable change in respiration, and the dog occasionally licks his chops. In 1910, Pavlov (5) termed this directional movement the orienting reflex (*OR*). At first, he was only concerned with controlling the *OR*, since it interfered with the conditional response. The sound-shielded conditioning chamber was constructed to reduce the interference of the extraneous stimuli invariably present in a conditioning laboratory: a sudden gust of air, a speck of plaster falling from the ceiling, accidental noises, etc. Pavlov gradually came to see, however, that the *OR* was an important bit of behavior in itself, and in 1916, he discussed inquisitiveness as an elaboration of the *OR* (5).

In 1947, Robinson and Gantt (6) reported that the *OR* in dogs has both motor and autonomic components. They studied the changes in heart rate (*HR*), respiration, and salivation in addition to the directed motor movement.

---

* Received in the Editorial Office on April 25, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This work was done in the Pavlovian Laboratory of the Johns Hopkins Medical School.

[2] This investigation was supported in part by a research grant (M-1091) from the National Institutes of Mental Health, Public Health Service, and by a grant from the American Heart Association.

Most stimuli that are to be used as conditional stimuli elicit the *OR*, and it is generally desirable to extinguish the *OR* before initiating the reinforcement procedure. In the *OR* phase of the autonomic study mentioned above, the dogs were given 5 to 15 repetitions of each of three nonreinforced tones/day for 20 to 30 daily sessions. The tones were 256, 512, and 1024 dvs, and the interval between tones was two minutes. Each tone lasted five seconds, and the order of presentation was always the same, T-256, T-512, and T-1024.

The purpose of the study was to compare the rate of formation of cardiac and motor conditioned responses, and it was desirable to approximate the intensities of the unconditioned stimuli prior to the reinforcement procedure (2). During the middle of the orienting training, several electrical stimuli were administered to the dogs at the end of two daily *OR* sessions (following the tones). All the dogs were to receive moderate electrical stimuli.

As a result of difficulties in the electrical circuit, Schnapps was accidentally given on the first day of the approximation procedure three electrical stimuli of a very high intensity (120 volts, 60 cycle, a.c.). Each stimulus lasted from two to five seconds, and the interval between stimuli was approximately five minutes. While the stimuli were applied to the left forepaw, the current radiated to the cardiac electrodes. These electrodes were located on the two sides of the animal's body, one anterior and the other posterior to the heart.

Schnapps had received 253 nonreinforced repetitions of the three tones prior to the session in which the electrical stimuli were administered. His behavior was relatively unstable during the early *OR* sessions as compared with the other animals. He alternated through successive periods of excitability and somnolence. But the *OR* had largely extinguished (both the motor and cardiac components) prior to the approximation procedure. The behavior of Schnapps to the electrical stimuli and in the sessions after these stimuli is summarized below.

Each electrical stimulus produced an immediate withdrawal reaction followed by struggling. The duration of the struggling increased with the successive stimuli, and the animal struggled for about five minutes following the third shock. The experimenter (*E*) entered the conditioning chamber after each stimulus attempting to determine the source of difficulty in the electrical circuit. Schnapps' struggling behavior disappeared at the sight of *E* after the first two shocks, but not after the third.

The first shock resulted in piloerection and an odor of feces. After the struggling subsided, the animal withdrew to the back of the conditioning stand. He held up his stimulated paw for several seconds. He again raised his paw to the second shock and held it up during the third shock and for several minutes thereafter, except during the bursts of struggling. Each beat of the heart was distinctly visible through the abdominal wall, and while *HR* was not recorded it was estimated to be about 190 to 200 beats/minute. The animal urinated and defecated to the second shock and vomited to the third shock. A sexual erection was noted following the third shock. For the first time, the dog refused to eat on leaving the conditioning stand.

On the day following the electrical stimuli, Schnapps was more restless than usual and his *HR* was accelerated (see Figure 1). The dog again refused to eat following the *OR* procedure. His behavior was interpreted as a mild upset, and the striking changes that followed were not expected.

During the next 24 days Schnapps' behavior became more pathological: (*a*) Heart rate in the experimental room increased during the next few days reaching a maximum of about 153 beats/minute. (*b*) Restlessness gradually increased and gave way to struggling on the 11th post-shock day. (*c*) A muscular tremor developed in the upper muscles of the shocked foreleg on the 12th post-shock day. (*d*) On the 14th post-shock day the animal began to hold the stimulated forepaw up for variable periods of time. (*e*) On the 18th post-shock day the animal for the first time exhibited signs of fear on seeing *E* and the animal caretaker. He crouched in the corner of the kennel with his tail between his legs, and he was hesitant to leave the kennels. (*f*) The animal refused to come to the experimental room on the 22nd day by way of the stairs, but showed no hesitance about the coming by elevator. Normally, the animals were brought to the experimental room by way of the elevator; but when the elevator was busy, the stairs were used. (*g*) Urination, defecation, vomiting, and sexual erections appeared on the 24th day.

The above behaviors intensified once they appeared, and when a new behavior appeared it persisted. Within any one daily session the order of appearance of the behaviors approximately paralleled the order in which they emerged on a day-to-day basis. It is to be noted that certain new reactions not present at the time of the electrical stimuli developed: fear of *E*, the animal keeper, and the stairs.

E's concern throughout all of this was that of attempting to salvage the dog for conditioning experiments. During the 24-day period during which these pathological symptoms appeared and intensified, numerous attempts were made to extinguish the pathological behavior. Special food and care were given, multiple daily sessions were run, and the experiments were conducted with the door to the conditioning room open.

The animal was moved to another conditioning chamber on the 21st post-shock day, and the usual orienting session was conducted by other staff members. The motor and cardiac components of the *OR* were exaggerated in this novel room, as compared with pre-shock behavior, but the dog behaved well, and his *HR* was appreciably lower than on the previous day (see Figure 1). On the 24th post-shock day, white lights (10 repetitions) were substituted for the usual tones, and this resulted in an even greater intensification of the previously observed pathological behavior. Schnapps vomited six times, defecated twice, urinated several times, and exhibited several sexual erections. Vomiting and sexual erections were also observed during the next two-day period when the usual *OR* procedure was resumed.

At this point, "extinction" was attempted by giving a series of mild electrical shocks. On the first day of this procedure, the animal was given 15 mild stimuli at one-minute intervals. The stimuli were regulated to produce a consistent but barely observable flexion response. The animal became remarkably quiescent after the first five or six stimuli and ate several dog biscuits after leaving the conditioning stand. This was the first time he had eaten in the experimental room since the traumatic shocks.

On the next day, 20 mild stimuli were given. Struggling had disappeared and the *HR* was decreased. Schnapps did not attempt to avoid *E* in the kennel this day. The two-day procedure of mild stimulation, which Whitehorn (7) has called a reassurance of mildness, completely reversed within one week all the pathological behavior except the animal's hesitation about coming to the experimental room by way of the stairs. This latter behavior persisted until his death about a year later. Schnapps was sufficiently improved that he could be used in the conditioning phase of the investigation; and his conditioning records were comparable to the other animals except that his *HR* was somewhat higher during the first 200 trials (1).

Figure 1 presents the mean *HR*'s for Schnapps in various sessions preceding and following the electrical stimuli. The *HR*'s to the three tones were pooled and a trial refers to one repetition of the three tones. It may be seen that the cardiac component of the *OR* was increased following the traumatic stimulation for a number of sessions. This suggests that the *OR* may serve as an indicator of the emotional state, and this hypothesis has been recently tested in humans (3).

The electrical circuit was repaired before the other three animals were stimulated, and while they received shocks of a sufficient intensity to cause them to whine, the intensities of these shocks were a fraction of the intensity received by Schnapps. This procedure had the effect of increasing the *HR* of these dogs to the tones, but this was restricted to the first session following the electrical stimulation. The extinction of the *OR* continued with only a transitory interruption.

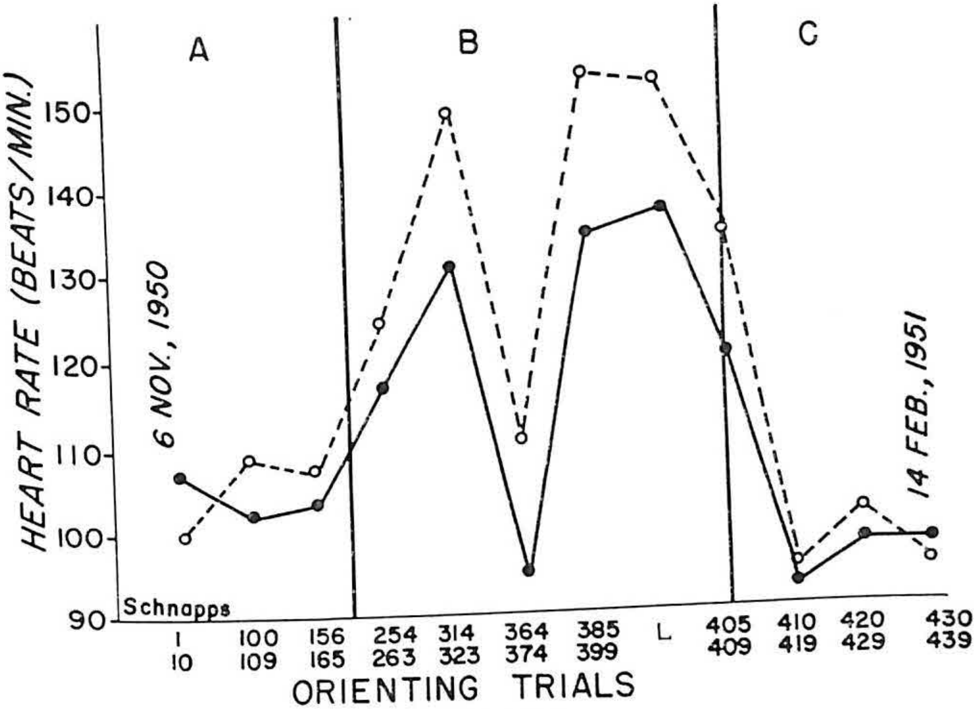

FIGURE 1

Mean *HR*'s for various sessions preceding (A) and following (B) the traumatic electrical stimuli and for various sessions during and following the mild electrical stimuli (C). For this presentation the *HR*'s to the three tones were pooled. The solid line refers to the five-second prestimulus period and the dotted line to the five-second stimulus period. The abscissa indicates the specific trials from which the prestimulus and stimulus means are derived. Trials 254 - 263 occurred on the first postshock day; Trials 364 - 374 were given in a different room; *L* refers to a session in which a white light was substituted for the usual tones; and Trials 405 - 409 were given after the last series of mild electrical stimuli, on the same day.

The development of pathological behavior such as has been described for Schnapps is termed autokinesis (2, 4). This refers to the spontaneous evolution of new behavior in the absence of further "prepotent" stimulation; in the present case, the electrical stimuli. It has been theorized that autokinesis may be maladaptive or adaptive as in the pathological behavior of Schnapps following the traumatic stimuli and his adjustment following the mild stimuli.

SUMMARY

This report has described the pathological behavior of one dog as produced by traumatic electrical stimuli and his subsequent recovery as a result of mild or reassuring shocks. Three other dogs given stimuli lesser in intensity failed to develop pathological symptoms. Data on the cardiac component of the orienting response have been presented, and these data suggest that this response is exaggerated by emotional upsets.

REFERENCES

1. DYKMAN, R. A., & GANTT, W. H. Relation of experimental tachycardia to amplitude of motor activity and intensity of motivating stimulus. *Amer. J. Physiol.*, 1956, **185**, 495-498.
2. DYKMAN, R. A., GANTT, W. H., & WHITEHORN, J. C. Conditioning as emotional sensitization and differentiation. *Psychol. Monog.*, 1956, **70**, 1-17.
3. DYKMAN, R. A., REESE, W. C., GALBRECHT, C. R., & THOMASSON, P. J. Psychophysiological reactions to novel stimuli: Measurement, adaptation, and relationship of psychological and physiological variables in the normal human. *Ann. N. Y. Acad. Sciences*, 1959, **79**, 43-107.
4. GANTT, W. H. Principles of nervous breakdown—schizokinesis and autokinesis. *Ann. N. Y. Acad. Sciences*, 1953, **56**, 143-163.
5. PAVLOV, I. P. Lectures on Conditioned Reflexes. (Trans. by W. H. Gantt.) New York: International Publishers, 1928.
6. ROBINSON, J., & GANTT, W. H. The orienting reflex (questioning reaction): cardiac, respiratory, salivary, and motor components. *Bull. Johns Hopkins Hosp.*, 1947, **80**, 231-253.
7. WHITEHORN, J. C. Introduction and survey of the problems of stress. In *Symposium on Stress*. Washington, D. C.: Army Medical Center Graduate School, Walter Reed Army Medical Center, 1953.

University of Arkansas
Medical Center
Little Rock, Arkansas

Johns Hopkins Medical School
Baltimore, Maryland

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 111-117.

INTERVIEWING CHILDREN IN A SCHOOL SETTING*

Department of Psychology, Wheaton College

DAVID ELKIND

A. INTRODUCTION

Interviewing children is now a generally accepted and widely used tool in child development research. As yet, however, there has been little discussion of the general principles involved in conducting a successful child interview. The rules given by Piaget (1929) deal primarily with the interpretation of responses and not with the interpersonal relation between child and adult. Discussions such as that of Goodenough (1950) touch on only some of the problems of child interviewing, namely those involved in diagnostic testing within a clinic setting. The principles given in the present paper are directed primarily towards the research psychologist but are general enough to be applied by other psychologists and professionals who work in a school setting.

The need for competence in conducting the child interview goes beyond the desire for reliable data and valid measurement. The research psychologist working in the local school is taken by the community as the prototype of all psychologists. Whatever impression he leaves will often determine the reaction of the community to future researchers, clinicians, and to psychology generally. Because teachers have their own needs and problems the psychologist will inevitably provoke some resentments and hostilities. These resentments and hostilities, however, can be minimized by the exercise of tact, courtesy, and common sense.

B. ADVANTAGES OF THE SCHOOL SETTING

Despite the greater responsibility demanded of the psychologist doing research in the schools, he has many compensating advantages. Perhaps the greatest advantage is that he sees children in their natural setting. At a nursery or public school the child is surrounded by familiar people, sights, and sounds. The child feels at home and displays a spontaneity and freedom that may never appear in the clinic setting. Another advantage is that the child is away from home and the center of attention. The youngster's motivation to be alone with an adult is much greater than it would be if the child's

---

* Received in the Editorial Office on April 26, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

parents were nearby. Finally, from a research point of view, the psychologist working in a school has the advantage of being able to see a large number of youngsters in a short period of time. These advantages easily balance the drawbacks to working in the public schools.

C. Relations with the School

An excellent set of suggestions for psychologists working in the schools is given by Ira Iscoe (1959). The rules given by Iscoe cover the major obligations of the research worker to the teacher and school officials. There are, however, several minor courtesies the psychologist would be wise to perform. The teachers and the principal's secretary should be told of the days and hours when the psychologist will be at the school. If for any reason he cannot be at school during his scheduled time, then it is incumbent on him to notify the school as far in advance as possible. The psychologist should always let the principal's secretary know of his arrival and of his departure. Finally the psychologist should acquaint himself with the school's regulations regarding smoking, fire drills, assemblies, and he should conduct himself and his research activities accordingly.

D. The Psychologist's Introduction

How the psychologist is introduced depends upon the age of the children involved. Before the age of about six years many children are still wary of strangers and need more time than older children to adapt to new adults. When he works with the nursery school and kindergarten age groups the psychologist must, therefore, plan on spending several days as a passive observer in the group. And he must be *passive!* He should take no active part in the routines of the group unless asked to do so by the teacher. If children approach him, he is wise to answer their questions simply and without trying to strike up an acquaintance or to engage in lengthy conversation. The psychologist should also be careful not to help children on with their wraps, with tying their shoes, or with any other skill they are attempting to master.

Whenever the psychologist takes an active role in the group he puts himself in competition with the teacher's authority. The children whom he has helped will look upon him as a special protector which lessens the teacher's control. Furthermore the help itself is often unnecessary and undesirable. Many children will "con" the unwitting and idle adult into performing a feat of which they themselves are quite capable. If the psychologist gives in to children's pressures he runs the risk of injuring the relationship with the teacher whose cooperation he badly needs.

Most children past the age of six years have already had considerable school experience and on the whole adapt easily to new people. These children can generally be interviewed immediately after the psychologist has been given a simple introduction like "Mr. *X* works at *Y* and he has come to play games (or to talk) with you." Even if the psychologist has a Ph.D. he is wise to be introduced as Mr. and so avoid the unpleasant associations often connected in the child's mind with the word, "doctor."

Regardless of what age group he works with, the psychologist's appearance is of great importance. Children are much more sensitive to colors, sounds, and smells than are adults. Lacking experience, they are likely to react to the adult on the basis of whether or not he is aesthetically pleasing. A psychologist who is untidy, speaks harshly or smells strongly of tobacco is likely to set up resistances in the child which will lessen the chances for a successful interview.

E. The Interviewing Room

Working in public or nursery schools means that one must make use of borrowed quarters. The psychologist should make whatever room he uses as pleasant as possible by letting in plenty of fresh air and light. It is a good idea to have some toys in evidence to make the children feel more at home. The only essential furnishings, however, are a child sized table and chairs.

The psychologist will have more success if he uses a low chair which brings his eyes on a level with the child's. When the child is at the psychologist's eye level the size differences between child and adult are minimized. Reducing the awareness of size differences lessens feelings of littleness and inferiority. It is also a good practice not to sit directly facing the child. A comfortable arrangement is to have the youngster at right angles to the adult. In this way the child is relieved of being under the constant scrutiny of the psychologist.

F. The Child's Introduction to the Interview

How one introduces the interview again depends upon the age group involved. The main problem with the younger children is to reduce their fear of being alone with the psychologist. The simplest and most direct approach is to ask the child his name and his age. These are facts the child knows well and giving them correctly builds self-confidence and reduces anxiety.

If the child has on or is carrying what is obviously a new or prized possession it is quite appropriate to remark "how pretty it is" or something similar. The compliment, when given sincerely, is an invaluable technique for building rapport. Sometimes questions about approaching holidays, birthdays, or school affairs can be used to convey the psychologist's friendly interest. Regardless of the technique the psychologist must make sure that self-confidence and rapport are present before he begins the structured part of the interview.

In contrast to the young children whose greatest fear is of the psychologist, the older children are most anxious about the interview itself. Testing has become so much a part of children's educational program that even first grade youngsters already show wariness and anxiety about being tested. In addition to test anxiety children may also be apprehensive over "stories" about the psychologist made up by mischievous youngsters (or unthinking teachers).

Apprehensiveness in the child is shown most directly in his posture. The tense anxious child sits on the edge of his chair or in an awkward, uncomfortable position. With these children and indeed with all children the psychologist's first task is to make sure that the child is physically comfortable. Once this is accomplished the interviewer can say "I guess you have heard about what I'm doing." This affords the child a chance to voice and to make light of his apprehensions like one boy who replied to the writer "Yeh, you are asking crazy, mixed-up questions." Such expression is often enough to relieve much of the child's tension. With overly excited children it is sometimes useful to have them take a deep breath in imitation of the psychologist. This usually suffices to calm them.

When the child is relaxed the psychologist should tell the child (only if it is true) that he is not there to test him, but only to find out what boys and girls think about certain things. That most children will not believe such a statement is beside the point. What the statement does is to show the child that the psychologist is aware of the child's anxieties and wants to reduce them. It works, therefore, towards building rapport.

The amount of time spent in the introduction should vary with the child. Some confident, lively, and gregarious children may require only a little preparation. On the other hand, shy, timid, and frightened children may need to be talked with for some time before the structured part of the interview can begin. The talk should center on some real interest of the child about which he can speak freely and well. Once he has built up his self-confidence in this way and is assured of the psychologist's friendly and noncritical interest, even the shy child will do his best to cooperate.

G. How to Proceed

To avoid repetitious answers and early boredom it is well to have questions interestingly worded and arranged. Similar questions should be separated in time. If the child hesitates he can be encouraged with um hum's?, yes?, and?, anything else?, which prompt, without directing, his thought. The child should never be pressured to the point of arousing anxiety. If he does not respond to one or two promptings the question should be dropped.

One of the biggest causes of faulty data is to assume that the child knows the meaning of terms he uses correctly. The child, no less than the adult, understands more words than he uses, but also can use many more words than he understands. Whenever the child employs an abstract term he should be asked general questions like "How do you mean?", "Tell me more about it?". Answers to these questions usually suffice to indicate how well the child understands the term.

H. When to Stop

Anyone planning to interview children is wise to "keep it short." Unlike adults who can converse for hours, most children will grow restless after 10 or 15 minutes of straight talk. The longer the interview, the more varied the questions and technique must be if interest is to be maintained. There is seldom, however, a legitimate reason for extending an interview beyond 15 or 20 minutes.

Sometimes it is necessary to stop earlier than planned. There are several warning signs. One sign is the obvious answer at random and a second sign is the repetitious "yes" or "no" to everything the psychologist asks. Usually these answers arise from disinterest and boredom. There is little the psychologist can do at this point for when negativism is present it goes in only one direction—*up*. Since negativism is directed toward the psychologist and not the questions it is futile to attempt to regain interest by a change in the questioning. A sensible procedure is to permit the child to return to class and to see him on another day. In a good many cases the negativism will subside with time. The writer, however, has had the experience of a kindergarten child bursting into tears every time he appeared.

Another sign that should alert the psychologist is an attempt by the child to discourse on an unrelated subject. There are several possible reasons for this: (*a*) The child is tired or bored with the interview. If boredom is the cause, it is best to end the interview and see the child another time. (*b*) The child is overly excited by some enormous event in his life; a baby brother, a birthday party, etc. If the psychologist will listen, ask interested questions, and make a few appreciative comments he can usually continue after the child is through. (*c*) The child feels he is doing poorly and is trying to escape the situation. This is a more rare motivation and usually appears in children who show in other ways strong needs for achievement coupled with feelings of inferiority. Unfortunately there is little the research psychologist can do or say that will remedy a personality problem. He should assure the child briefly and simply that he is doing "all right." Any attempt at excessive assurance will be recognized for what it is and the child will feel confirmed in his conviction of inadequacy. The psychologist should complete the session with the child. A remark to the teacher about the problem is generally in order but should be tactfully made.

I. Rewards and Punishments

Rewarding children with candy has some justification in nursery schools and in schools located in blighted areas. Unless, however, candy is called for by the design of the experiment it is best avoided. If only one or several children in a classroom have candy, this can be a cause of disturbance and may alienate the teacher. More importantly the use of candy may lull the psychologist into relaxing his efforts to establish and maintain rapport because he feels cooperation is assured. Genuine cooperation, however, is never won with a bribe.

Bribes are not necessary for the properly conducted interview has its own rewards for the child. One of the child's rewards is being alone with an adult who is interested in, and even writes down, what he says. To most children this is a rare experience and they enjoy the sense of power given by dictating to an adult. Only a few children will carry it to the extreme of an officious youngster in the writer's experience who repeatedly ordered "strike that out." Visibly writing down what the child says is an important reward to the youngster and should not be supplanted by tape recordings.

Another type of reward should be given with decreasing frequency at advancing age levels. Verbal rewards like "Very nice," "Just right," etc., can be given liberally almost after every response at the nursery school and kindergarten ages. After ages 5-6 children are more sophisticated and too frequent verbal rewards will be dismissed as insincere or will be anxiously watched as a delicate barometer of successful performance. Generally questions fall into units or sets and verbal rewards should be given after each set. One need not be hypocritical for there is always something nice one can honestly say about a child's performance.

It is not the responsibility nor the prerogative of the psychologist to punish a child who does not cooperate or for any other reason. The child is rightly thought of as the psychologist's guest and should be treated as such. If the child misbehaves he should be sent back to the classroom without reprimand. Like any other unpleasant guest he is not abused but simply not asked to return. For most children this is punishment enough. Whether the child has cooperated fully or not at all, he should always be thanked at the close of an interview.

J. QUALITIES OF THE PSYCHOLOGIST INTERVIEWING CHILDREN

Talking successfully with children is a skill. Like any skill some people pick it up almost without effort, others learn it with practice, while still others can never acquire it. Some of the qualities demanded of a person interviewing children are a genuine liking and respect for children, and an awareness of their hopes and fears. He must also have a sense of humor broad enough to see the often irrational and comic quality of adult behavior in general and of his own behavior in particular.

In acquiring the skill there is no substitute for practical experience. Happily, schools, day camps, and summer camps are more open now to psychological investigations. A psychologist who wishes to study children has resources, therefore, that were not available 20 years ago. He also has new responsibilities. The psychologist working in the schools represents his entire profession. If he puts his personal relationships with teachers and children ahead of his data collecting zeal, then both his own and the community's experience is likely to be a happy one.

REFERENCES

1. GOODENOUGH, F. L. Mental Testing. New York: Rinehart, 1950.
2. ISCOE, I. On harmonious relations with public schools. *Amer. Psychol.,* 1959, **14**, 694-5.
3. PIAGET, J. The Child's Conception of the World. London: Kegan Paul, 1929 (reprinted in 1951).

Department of Psychology
Wheaton College
Norton, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 119-131.

PERCEIVED FRONTAL SIZE AS A DETERMINER OF PERCEIVED BINOCULAR DEPTH*

US Army Medical Research Laboratory, Fort Knox, Kentucky

WALTER C. GOGEL[1]

A. INTRODUCTION AND PROBLEM

When two objects, one of which is more distant than the other, are viewed with both eyes, a binocular retinal disparity is produced. This is illustrated in the top view schematic drawing shown in Figure 1. In this figure two objects (*e* and *f*) are presented at different distances from the *S*. The retinas of the left and right eye of *S* are shown as R_l and R_r respectively. The distance "A" represents the distance between the nodal points (N_l and N_r) of the two eyes. The physical distance between Objects *e* and *f* is shown as X_{ef}. The retinal distance between Objects *e* and *f* is P_lQ_l in the left eye and P_rQ_r in the right eye. The difference between these retinal distances is the binocular disparity between the two objects. It can be seen that the binocular retinal disparity is proportional to $a_e - a_f$ and therefore $a_e - a_f$ can be used as a measure of binocular disparity.

As a consequence of the binocular disparity between Objects *e* and *f*, Object *f* will be perceived to be behind Object *e*. The problem of this paper is to investigate the process which determines how far Object *f* will appear behind Object *e*. It might be considered that the perceived depth between Objects *e* and *f* depends upon the distance which one or both of these objects appears to be from *S* (1, 6). From this point of view, a binocular disparity can be translated into a definite perceived interval only when a perception of a distance from *S* (a perceived absolute distance) occurs and, this perceived depth interval changes with the perceived absolute distance. Contrary to this point of view, the assertion is made and tested in this study that a perception of absolute distance is not required in order for *S* to translate $a_e - a_f$ into a perceived depth extent. It is asserted that the process of translating a binocular disparity to a perceived depth extent involves the perceptual interaction of frontal and depth extents when these are spacially adjacent in the visual field. By this it is meant that the depth extent represented by $a_e - a_f$ is necessarily

---

* Received in the Editorial Office on April 27, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank Jack A. Schneps for his help in collecting and analyzing the data of this study.

perceived in relation to whatever frontal extent is in its immediate vicinity. This process will be discussed with the aid of Figure 1.

The dotted line *ab* in Figure 1 represents an object of horizontal angular extent θ. Assuming that D_e is large compared with "A" (as is usually the case), the angle θ is also an angular measure of the retinal size of *ab*. In Figure 1, the frontal extent *ab* and the depth extent X_{ef} are in the depth and directional vicinity of each other. For purposes of discussion, these two extents together can be considered as representing a top view of a three dimensional object. It is the contention of this paper that the ratio of the perceived depth component to the perceived frontal width (or height) of this three dimensional object is proportional to the ratio of angular, binocular disparity to angular, frontal extent. If $\frac{1}{C}$ is the constant of proportionality in this relation, then this paper asserts that

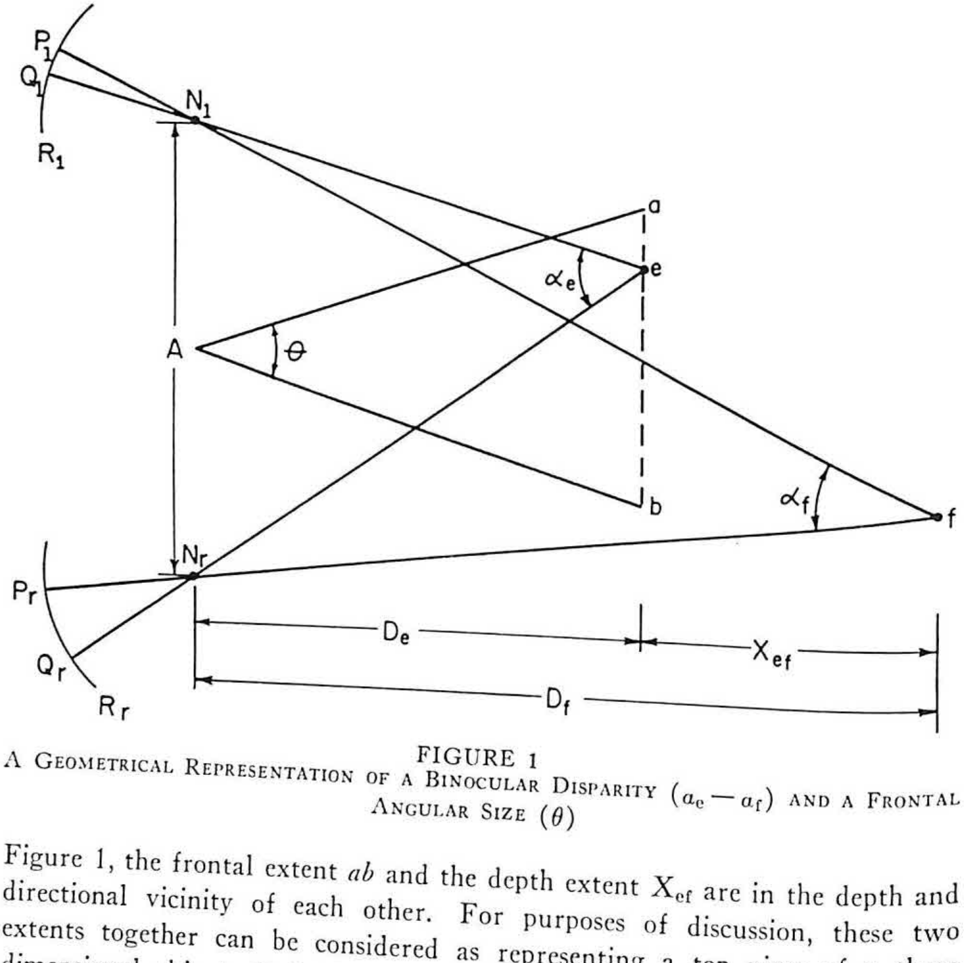

FIGURE 1
A Geometrical Representation of a Binocular Disparity $(a_e - a_f)$ and a Frontal Angular Size (θ)

$$m = \frac{\alpha_e - \alpha_f}{C\,\theta} = \frac{(\alpha_e - \alpha_f)'}{E} \tag{1}$$

where $(\alpha_e - \alpha_f)'$ represents the perceived extent associated with the binocular disparity $\alpha_e - \alpha_f$; where "E" represents the perceived extent associated with the frontal extent *ab*; and where "m" represents the perceived ratio of the depth extent to the frontal extent.

Equation 1 makes two assertions. It states that the perceived multiple (m) of X_{ef} to *ab* is determined only by $\alpha_e - \alpha_f$, θ and the magnitude of the subject constant *C*. Also, it states that by means of this perceived multiple, the size *E* which is perceptually associated with the frontal extent (*ab*) specifies the size which will be perceptually associated with the depth interval (X_{ef}). For example, if, in Figure 1, X_{ef} is perceived to be equal to *ab* ($m = 1$), then the perceived extent associated with X_{ef} will be equal to whatever size *ab* is seen to be. In this case, if the perceived size of *ab* is doubled, the perceived extent of X_{ef} will double, even though $\alpha_e - \alpha_f$ and θ remain unchanged.

This example could be used as a basis for experimentally testing Equation 1. However, it is more convenient to base the test on a slightly different situation. Equation 1 can be rewritten in the form

$$(\alpha_e - \alpha_f)' = \frac{E\,(\alpha_e - \alpha_f)}{\theta\,C} \tag{2}$$

The present study, consistent with Equation 2, will attempt to vary $(\alpha_e - \alpha_f)'$ by varying θ with both *E* and $\alpha_e - \alpha_f$ held constant. This means, in Figure 1, that neither the perceived size *E* of *ab* nor the binocular disparity $\alpha_e - \alpha_f$ between Objects *e* and *f* will be changed. However, the angular size θ of *ab* will be changed and the resulting change in perceived depth $(\alpha_e - \alpha_f)'$ will be measured. According to Equation 2, $(\alpha_e - \alpha_f)'$ should vary inversely with θ.

It will be noted that Equations 1 and 2 do not contain any term for perceived absolute distance. In fact, all the terms in the equations are considered to be independent of perceived absolute distance except possibly the term *E*. It might be expected that, for a constant value of θ, the perceived size *E* of *ab* would vary as some function of the perceived absolute distance [see Gruber (4) for a discussion of this problem]. But, perceived absolute distance is not the only factor which might be expected to determine perceived frontal size. Another factor is that of object familiarity. The perceived size of objects can differ or appear the same depending upon what they appear to be. For example, a rectangularly shaped object of constant angular size θ might appear to be small or large depending upon whether it is seen to be a postage stamp or a billboard. Or, because of unchanging familiar characteristics, the object might appear to be always the same size in spite of changes in its angular size θ. In the present study the latter situation will be used. The perceived size E will be kept constant in spite of changes in θ by having *ab* always look like the same familiar object.

It might be thought that changing the angular size of an otherwise identical familiar object would change its apparent distance. If this occurred, it would be difficult to distinguish between predictions made from Equation 2 and those made from the viewpoint which considers $(a_e - a_f)'$ to be dependent upon perceived absolute distance. However, an experiment which has been directly concerned with this possibility has demonstrated that the retinal size of a familiar object can be changed without inducing a change in its perceived absolute distance (2). This presents the possibility that the relation between $(a_e - a_f)'$ and θ in Equation 2 can be tested with the other terms of the equation held constant and that this can be accomplished without modifying perceived absolute distance.

B. METHOD

Some aspects of the task with which *S* is confronted in this study can be illustrated by Figure 1. Let Object *ab* and Object *f* be binocularly presented in an otherwise dark room at distances D_e and D_f from *S*. Suppose the Object *ab* is a normal-sized playing card while Object *f*, a ring of light, has no familiar characteristics to suggest any perceived size. Suppose also that when *S* looks at these objects the depth between them is perceived as being 10 times as great as the width of the playing card ($m = 10$). From Equation 2, if the perceived width (E) of the playing card is 2¼ inches, the perceived depth between the binocular objects should be interpreted as being 10 times 2¼ inches or approximately two feet. If, however, Object *ab* were a double-sized rather than a normal-sized playing card at distance D_e, the angular size θ of *ab* would be doubled. But, the perceived size of the playing card would remain unchanged and consequently the depth between Object *ab* and *f* would be interpreted as being 5 times 2¼ inches or approximately one foot. Equation 2 predicts that the perceived extent of the depth interval between the two objects should be less when the retinal size of the playing card is greater.

The results from an experiment illustrated by the above example must not only satisfy Equation 2; they must also distinguish between the process of perceiving a depth interval described by Equation 2 and a process dependent upon perceived absolute distance. For this reason, the change in the perceived distance between the objects as a function of a change in θ should occur without a change in the perceived absolute distance to the two sizes of playing cards. Finally, an additional requirement must be placed upon the experiment. This requirement is that only binocular cues should be available to determine the perceived depth *between* the objects. Presenting the playing card and the ring of light in a room in which no other objects are visible meets this requirement.

C. APPARATUS

1. *The Alleys*

A method was required (*a*) to measure the apparent distance between the cards and ring both when normal-sized and when double-sized cards were used and (*b*) to measure any change in the apparent absolute distance of the cards when the size of the cards was changed. The method used was the same as that employed in the previous study (2) in which the subjects threw darts to the apparent distance of the objects. The apparatus consisted of a long room divided into a viewing alley in which the cards and ring were located, and an adjacent and parallel throwing alley along which the thrown darts landed. A black curtain divided the throwing alley from the viewing alley so that the subject could not observe the results of his throwing. The front of the viewing alley was black and contained an aperture which permitted the subject to see into the viewing alley. The *S*'s position in front of the viewing aperture was only slightly to the left of the throwing alley. With this arrangement, the subject could throw darts down the throwing alley while looking into the viewing alley.

The floor of the viewing and throwing alleys consisted of the tops of a series of tables. The floor of the throwing alley was 9¾ inches above that of the viewing alley and was at the height of the viewing aperture. The bottoms of the cards and ring were always at the height of the floor of the throwing alley. A light-adaptation surface (of a brightness of 25 ft-1) was located to the left of *S*. By turning his head, *S* could look directly into the adaptation surface.

The experimental room was dark except for the cards and ring. The sides of the viewing alley were black, and the floor of the viewing alley was covered by a black cloth. In addition to the viewing aperture, a second, larger aperture was placed at a distance of 4¼ feet from the *S*. These precautions permitted illuminated objects to be presented in the viewing alley without causing any portions of the walls or floor of the viewing alley to be visible. The *S* could not see either of the two apertures when looking into the viewing alley.

The floor of the throwing alley was formed of soft fiber board into which the point of the darts could readily penetrate. The distance of the point of impact of the thrown darts from *S* was measured to the nearest centimeter. The darts were approximately 7½ inches long and weighed approximately 16 grams.

2. *The Stimuli*

The playing cards were photographic transparent positives of a seven-of-spades playing card. They were mounted in front of a light box and transilluminated with white light. When presented in a dark room they looked like illuminated playing cards and will be referred to as playing cards throughout this study. Figure 2 is a schematic drawing of the relative positions of the cards and ring. Two identical playing cards of the same physical size were presented simultaneously at the same distance from the subject (10 feet) with their inner edges separated by 5⅞ inches. Thus two playing cards representative of Object *ab* in Figure 1 were placed at the same distance rather than the single playing card mentioned in the previous example. A luminous ring (2 inches outer and 1¾ inches inner diameter) was placed

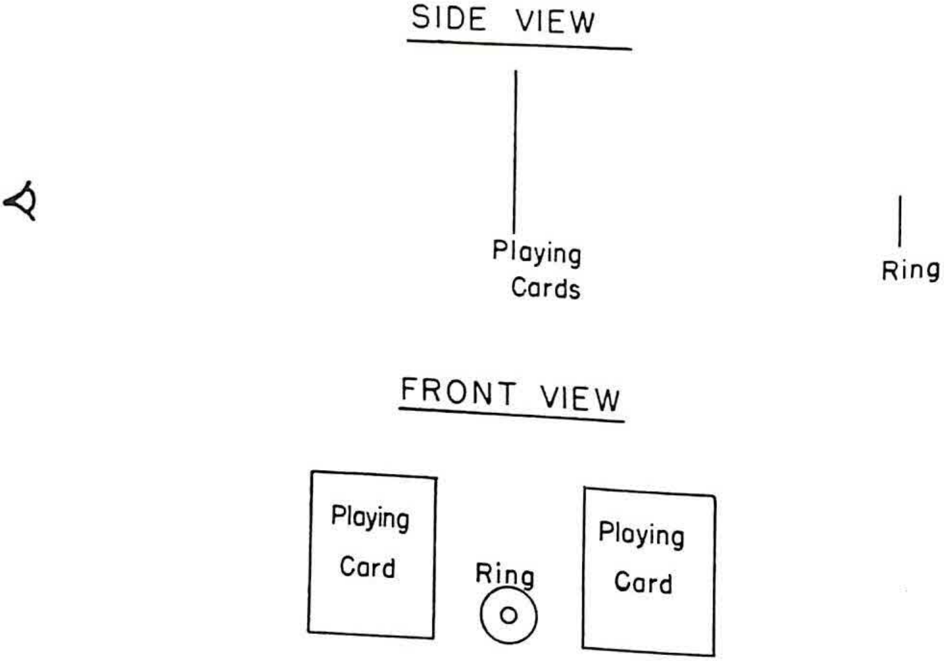

FIGURE 2

SCHEMATIC REPRESENTATION OF THE DISPLAY USED IN THE EXPERIMENTAL SITUATIONS five feet behind the playing cards. This ring was a transparent figure cut in an opaque black surface and mounted in front of a light box which transilluminated the ring with white light. The two playing cards were familiar objects. The ring was unfamiliar in the sense that its configurational characteristics were not such as to perceptually specify a particular size. The brightness of both the ring and the white portions of the playing cards was .39 ft-1. The two playing cards were either normal-sized ($2\frac{1}{4}$ by $3\frac{1}{2}$ inches) or double-sized ($4\frac{1}{2}$ by 7 inches) seven-of-spades playing cards. Two experimental situations were used, one consisting of the ring and the normal-sized cards, and the other of the ring and the double-sized cards.

3. *Subjects*

Two groups of 18 men were used as *S*s. They all were right-handed and they all had a visual acuity of at least 210 sec. of arc (as determined by using the Diagnostic Series for stereopsis of the Keystone Telebinocular). With one group of 18 *S*s, normal-sized cards were used and with the other group of 18 *S*s, double-sized cards were used. This was the only difference between the stimulus conditions for the two groups. With both groups, the playing cards were at a distance of 10 feet and the ring of light at a distance of 15 feet. The *S*s were randomly assigned to the two groups. The average interpupillary distance was 6.53 cm. and 6.46 cm. respectively for the group using the normal-sized cards and for the group using the double-sized cards.

D. PROCEDURE

1. *Experimental Situation*

Before participating in the experiment, each *S* was given 20 practice throws with the darts to accustom himself to the throwing task. This did not involve seeing either the playing cards or ring. When this was completed all the lights in the room except that from the card and ring were turned off, and a shutter covering the viewing aperture was raised. The *S* identified the playing cards and ring. He indicated that the ring was located behind the playing cards and that the two playing cards were equidistant from himself. The *S* was informed that he was to throw the darts so that they would land at the same distance from himself as the cards (a second set of throws was later made to the apparent distance of the ring). It was explained that the floor of the throwing alley was at the same height as the bottom of the cards and ring. The significance of this for his throwing was carefully explained. The shutter covering the viewing aperture was closed, and *S* was light-adapted by looking into the light adaptation surface for approximately five seconds.

The *S* then was handed a dart and he positioned his head with respect to the aperture. The light adaptation surface was turned off, the shutter was raised, and while looking in the viewing alley, *S* threw the dart down the throwing alley to the apparent distance of the cards. The place at which the dart landed was not visible to *S*. This process of light-adaptation and throwing the dart was repeated until five scoreable throws were completed with a scoreable throw defined as a throw in which a dart imbedded itself into the floor of the throwing alley. Following the completion of five scoreable throws to the apparent distance of the cards, the ring was designated as the target and five additional scoreable throws were completed. The *S* was then asked to estimate the distance of the cards and the ring from himself in feet. Binocular observation was used throughout the study.

There are several aspects of this procedure which require discussion. The reason for using two groups of *S*s rather than using all the *S*s in both of the experimental situations is related to the problem of determining whether the differently sized playing cards were perceived at different absolute distances. A previous study had indicated that different groups of *S*s must be presented with the different retinal sizes of the playing cards if the perception of the absolute distance of the cards is to be investigated (2). If the same group of *S*s had been used with both the normal and the double-sized playing cards, the results obtained from the throwing to the secondly presented playing cards could not have been interpreted as indicating the perceived absolute distance of these cards. A similar caution applies to the order in which the apparent distance of the playing cards and ring was measured. Successive throwing by the same *S* to objects at different apparent distances results in a measure of the apparent absolute distance of the first object and the apparent relative distance between the first and second object. It does not result in a measure of the apparent absolute distance of both objects. Since the apparent absolute distance of the cards and the apparent relative distance between the cards and ring were the measurements which were desired, the playing cards always were designated as the first and the ring as the second target. A more thorough discussion of the necessity for these procedures can be found elsewhere (2, 3).

2. *Calibration Situation*

The question of this study is whether the perceptual interrelation of frontal and depth extents can convert a binocular disparity to a perceived depth extent independently of perceived absolute distance. To answer this question, perceived distance scores, not throwing scores or verbal estimates are desired. For this reason, data from a calibration situation were used to transform the throwing scores and verbal reports which had been obtained in the experimental situations to perceived distance scores. In the calibration situation the *S*s threw darts and made verbal estimates under conditions in which an abundance of distance cues were available. Under these conditions, the obtained differences between *S*s in either verbal reports or throwing scores for a particular target would tend to be a consequence of differences in throwing and estimational habits and not a consequence of differences in perceived distance. From the results obtained in the calibration situation, individual calibration equations were derived, which when applied to the scores from the experimental situations transformed these scores to perceived distance scores. This transformation is discussed in more detail elsewhere (2, 3).

In the calibration situation, the overhead fluorescent lights were turned on in the viewing alley, the more distant restrictive aperture was removed and a checkerboard pattern of light and dark grey rectangles on the floor of the viewing alley was exposed. The sides and floor of the viewing alley were clearly visible during the calibration situation, but the throwing alley was in darkness. Instead of using the two playing cards and the ring, three white cardboard rectangles were simultaneously presented on stands at 10, 20, and 30 feet from the *S*. The bottoms of the rectangles were all at the height of the floor of the throwing alley. Approximately one hour after participating in an experimental situation, *S* was presented with the calibration situation. In the calibration, the *S*s made five scoreable throws to the apparent distance of each of the rectangles.

E. RESULTS

From the calibration situation, equations relating response scores to the physical distance of the rectangles were calculated for each *S*. This was accomplished by fitting a straight line by the method of least squares to the three data points resulting from the three target distances in the calibration situation. This resulted in two linear equations for each *S*, one equation from the dart throwing data and the other from the verbal estimates. Using these equations, response scores in the experimental situations were converted to perceived distance scores expressed in terms of physical distances in the calibration situation. Specifically, these perceived distance scores are the physical distances in the calibration situation which appeared to *S* to be equal to the distances he perceived in the experimental situation. The throwing scores or verbal reports from the experimental situations will be called "obtained scores," while the transformation of the obtained scores by the individual calibration equations will be called "perceived scores."

The average perceived and obtained distance scores from the experimental situations are shown in Tables 1 and 2 respectively. Each mean and standard deviation was computed from a distribution of 18 scores, one from each *S*. Since the interest of this study is in perceived distance, the results shown in Table 1 are particularly pertinent. The *t* test was used to determine the statistical significance of the various mean differences with the results shown in the table. From neither the verbal estimates nor the throwing (Table 1) was there any statistically reliable indication that the normal-sized cards were seen to be at a greater distance than the double-sized cards. But, from both of these types of perceived distance scores, the average perceived distance

TABLE 1

MEANS AND STANDARD DEVIATIONS OF THE PERCEIVED DISTANCE SCORES IN FEET FROM THE CARDS AND RING FOR TWO SIZES OF CARDS

Size of Cards	Cards Mean	Cards *SD*	Ring Mean	Ring *SD*	Mean diff.	*t* Value
Perceived Scores From the Throwing Data						
Normal-sized	13.9	4.0	27.9	7.6	14.0	10.27**
Double-sized	14.2	4.3	22.0	5.4	7.8	10.97**
Mean diff.	—.3		5.9		6.2	4.02**
t Value			2.61*			
Perceived Scores From the Verbal Reports						
Normal-sized	12.7	4.4	27.0	8.2	14.3	10.26**
Double-sized	11.3	5.8	23.0	15.1	11.7	4.88**
Mean diff.	1.4		4.0		2.6	.93
t Value	.80		.96			

* Significant beyond the .05 level.
** Significant beyond the .01 level.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF THE THROWING DATA AND VERBAL REPORTS IN FEET FROM THE CARDS AND RING FOR THE TWO SIZES OF CARDS

Size of Cards	Cards Mean	Cards *SD*	Ring Mean	Ring *SD*	Mean diff.	*t* Value
Throwing Data						
Normal-sized	10.2	2.8	18.7	5.1	8.5	10.33**
Double-sized	10.1	3.2	15.5	4.7	5.4	8.79**
Mean diff.	.1		3.2		3.1	3.05**
t Value	.10		1.93*			
Verbal Reports						
Normal-sized	10.0	5.6	20.7	12.1	10.7	5.83**
Double-sized	8.9	4.1	18.6	12.2	9.7	4.70**
Mean diff.	1.1		2.1		1.0	.38
t Value	.65		.51			

* Significant beyond the .05 level (one-tailed test).
** Significant beyond the .01 level.

between the cards and ring was greater with the normal- than with the double-sized cards. While this latter increase, as determined from the throwing data (6.2 feet), was statistically significant, the corresponding increase of 2.6 feet, as determined from the verbal reports, was not significant at the .05 level. The perceived distance scores from the verbal reports were analyzed also by the Mann-Whitney *U* test (5, pp. 116-127). From this test it was concluded that the perceived distance between the normal-sized cards and the ring was significantly greater (one-tailed test) than the perceived distance between the double-sized cards and the ring beyond the .05 level ($U = 100.5$). However, the Mann-Whitney *U* test produced no statistical support for the conclusion that the perceived distance of the cards was greater when the normal, rather than the double-sized cards were used ($U = 123.5$). It can be concluded that the perceived distance between the ring and cards decreased as the retinal size of the cards increased and that this change cannot be attributed to a change in the apparent absolute distance of the cards.

F. DISCUSSION

The results of this experiment agree with Equation 2. As θ was increased, with $\alpha_e - \alpha_f$ and E held constant, $(\alpha_e - \alpha_f)'$ decreased with this decrease not attributable to a change in perceived absolute distance. The process described by Equation 2 provides a simple explanation for the perception of depth resulting from a binocular disparity. It states that the *S* perceives a binocular depth interval in terms of an adjacent frontal extent and that the perceived size of this frontal extent is then used as a yardstick in perceiving the depth interval. It is clear that other predictions based upon the point-of-view of this study need to be explored. Several relevant investigations will be presented in future reports.

One limitation of Equation 2 should be noted. On the right side of Equation 2 the ratio $\frac{E}{\theta}$ occurs. Since this is a ratio it can be expressed as the perceived size per unit of angular frontal size. This permits the perceptual process used in translating $\alpha_e - \alpha_f$ to $(\alpha_e - \alpha_f)'$ to be independent of any particular frontal extent. For example, in Figure 1, $\frac{E}{\theta}$ can be considered as applying to all parts of a frontal plane through *ab* as well as to *ab* itself. But, in many or most situations it would be expected that different frontal planes occurring at different distance positions within the interval X_{ef} will have different values of $\frac{E}{\theta}$. For instance, in cases of perfect frontal constancy, $\frac{E}{\theta}$ will increase as the distance from S increases. But, as was discussed previously, it is expected that the value of $\frac{E}{\theta}$ at a particular distance is used in determining the perception associated with an increment of binocular depth at that distance. This means that Equation 2 should be written as a differential equation, and therefore,

$$(a_e - a_f)' = \int_f^e da' = \frac{1}{C}\int_f^e \frac{E}{\theta}\, da \qquad (3)$$

where da' is the perceived increment of depth associated with the physical depth increment da. It will be noted that $\frac{E}{\theta}$ can vary within the interval $a_e - a_f$ in Equation 3; whereas in Equation 2, $\frac{E}{\theta}$ was assumed to be a constant throughout this interval. A method of evaluating Equation 3 will be considered in future reports.

G. Summary and Conclusions

This study is concerned with the process by which a binocular disparity is translated into a perceived depth interval. It is proposed that S will perceive a binocular depth interval to be some multiple of the frontal extent of an adjacent object. By means of this multiple the perceived size of the adjacent object then becomes the linear yardstick which translates the binocular disparity to a perceived extent. It is proposed that the perception of absolute distance is not necessary in order for this process to occur. In an experiment involving different frontal sizes of the same familiar objects, the results indicate that the perceived depth associated with a constant binocular disparity increased as the retinal size of the familiar objects decreased. This result cannot be explained by a change in the perceived absolute distance of the familiar objects but is consistent with the point of view of the study. The point of view of the study is expressed by an equation which predicts the apparent depth extent associated with a binocular disparity as a function of (*a*) a subject constant and (*b*) the ratio of the perceived to the retinal size of frontal extents in the vicinity of the binocular disparity.

References

1. Fry, G. A. Visual perception of space. *Am. J. Optom.*, 1950, **27**, 531-553.
2. Gogel, W. C., Hartman, B. O., & Harker, G. S. The retinal size of a familiar object as a determiner of apparent distance. *Psychol. Monog.*, 1957, **70**, 1-16.

3. ———. Perceived frontal size as a determiner of perceived stereoscopic depth. US Army Med. Res. Lab., Report No. 296, 1957.

4. Gruber, H. E. The relation of perceived size to perceived distance. *Amer. J. Psychol.,* 1954, **67**, 411-426.

5. Siegel, S. Nonparametric Statistics for the Behavorial Sciences. Toronto & London: McGraw-Hill, 1956.

6. Von Kries, J. Notes on perception of depth. *In* H. V. Helmholtz's *Treatise on Physiological Optics.* Vol. III (J. P. C. Southall, *Ed.*). Menasha, Wisconsin: Banta, 1925.

US Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 133-138.

THE BEHAVIOR OF PRISONERS IN A "PRISONER'S DILEMMA" GAME*

Departments of Psychology, Ohio State University and Case Institute of Technology

ALVIN SCODEL AND J. SAYER MINAS[1]

A. INTRODUCTION

Although many behavioral scientists are concerned with the applicability of formally derived models to behavior, the germinal work of von Neumann and Morgenstern (1947) on game theory has eventuated in very few empirical studies. Recent interest in decision-making processes can be attributed in part to the von Neumann and Morgenstern program, but the personality and social correlates of behavior in games, with the exception of a few scattered studies such as that of Deutsch (1958), remain largely unstudied.

The present investigation is an extension of a study that utilized college students as subjects. Both the initial study (Scodel, Minas, Ratoosh, and Lipetz, 1959) and the present one employed that version of a two-person non-zero-sum game known as a "prisoner's dilemma." This situation as discussed by Luce and Raiffa (1957) is characterized as follows: a district attorney who lacks sufficient evidence for conviction separates two suspects and informs each that he has two alternatives: to confess to the crime the police are sure the pair has committeed or not to confess. If neither confesses, the prisoners are informed that they will receive minor punishment. If both confess, a sentence less than the most severe one will be recommended for both. If one confesses and the other does not, the one who confesses will get off almost scot-free, but the other will receive the most severe sentence possible.

In our particular variation of the prisoner's dilemma monetary values were assigned to outcomes and the game was repeated for 50 trials. In the initial study it was found that college students, almost without exception,

---

* Received in the Editorial Office on May 10, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in whole by the United States Air Force under Contract No. AF 49(638)-317 monitored by the AF Office of Scientific Research of the Air Research and Development Command. Special thanks are due to Dr. Antonio Suziedelis and Warden Walter Jacquot of the Chillicothe Federal Reformatory for their aid in obtaining subjects for the study and to Dr. David Marlowe and Harve Rawson for their help in the collection of the data.

played a non-collaborative strategy. The behavior of actual prisoners in such a game has intrinsic appeal for several reasons. First, it affords a basis for comparing middle-class subjects with an anti-social group. Second, and more important, it offers an opportunity for testing the notion of "honor among thieves." A collaborative strategy would indicate that subjects were attempting to maximize joint return at the expense of the experimenter. Conversely, a non-collaborative strategy would suggest that the subjects were less concerned with the extraction of a mutually beneficial reward than they were with vying with each other.

B. METHOD

Eighteen pairs of subjects were obtained from the population of the Federal Reformatory at Chillicothe, Ohio. Of the 36 *S*s the large majority had been convicted on robbery charges involving government property. A few had been imprisoned for transporting stolen cars across a state line. All *S*s were between the ages of 17 and 25. In almost all instances each member of a pair of prisoners was familiar with the other by sight, but in no case were members of a pair "buddies." This was insured by selecting pairs whose members not only lived in different cell blocks, but had different work assignments. The only other criterion imposed for selection was a level of brightness sufficiently high to understand the game. Every inmate at the institution has taken a Stanford Achievement Test and an Army Beta test. Those whose scores were sufficiently low to merit doubt about their ability to understand the instructions were excluded from further consideration.

Members of pairs of *S*s were seated on opposite sides of a partition. Before each *S* was a panel containing a red and a black button. The following instructions were given by the experimenter, who sat at a right angle to the pair:

> Each of you will have a chance to make some cigarettes in this situation. This is what each of you will do. You see in front of you a red button and a black button. On each trial you will press either the black button or the red button. If you press the black button, two things can happen. If you press the black button and the other person also presses the black button, you get three cigarettes and he gets three cigarettes. If you press the black button and he presses the red button, you get nothing and he gets five cigarettes. Suppose you push the red button. Again, two things can happen. If you push the red button and the other person presses the black button, you get five cigarettes and he gets nothing. If you push red and the other person also pushes red, you get one cigarette and the other person gets one cigarette. You will keep all the cigarettes that I will give you.

Cigarettes from opened packs were used for payoffs in the game. However, a number of cartons of assorted standard brands were in clear view of the *S*s, and they were informed that at the end of the game they could obtain a pack of whichever brand they preferred for each twenty cigarettes accumulated in the game. Thus, an *S* who accumulated 107 cigarettes, for example, was given five packs in addition to seven cigarettes from an open pack. It should be stressed that cigarettes are the principal if not the only medium of exchange in the reformatory and therefore have considerable value, even for non-smokers.

The instructions were repeated for each pair so that there would be no misunderstanding about the payoff matrix. Moreover, the matrix was posted on the partition in front of each *S*, who was free to refer to it at any time during the game.

The game ran for 50 trials. When *S*s pressed their buttons, the combination of buttons which had been pressed was registered by a light on the examiner's panel, and *S*s were paid off in cigarettes after each trial. Although *S*s did not communicate with each other, the game contained no hidden information. If, for example, an *S* played red and received one cigarette, he knew that the other player had also played red; if he received five cigarettes, he knew that the other player had played black. In some cases the game was stopped short of 50 trials if each player pressed the same button on eight consecutive trials; it was assumed that the combination would be played for the remainder of the 50 trials.

The game was conducted in a large room with optimal privacy. All pairs were run on the same day so that game performance would not be affected by discussion among the prisoners. Inspection of the data revealed no discernible differences between the 12 pairs run in the morning and the six afternoon pairs.

C. RESULTS

In a sense the college and prison groups are not comparable since the payoff matrix for the college group, despite an identity in numbers, involved pennies rather than cigarettes, and it seems reasonable that one cigarette has more utility for a prison inmate than a penny has for a college student. The results for the two groups are presented only to demonstrate the over-all lack of collaboration in the two groups.

In 22 pairs of college *S*s run in this game, the red-red choice was the modal one in 17 pairs; black-black (collaboration) was the modal one in only two pairs. The median number of black-black plays was five whereas the median number of red-red plays was 21. The number of red-red plays exceeded the number of back-black plays in 20 of the 22 pairs. To put it another way, only two of the 22 pairs collaborated in the game.

In the prison population, which is the focus of the present study, the red-red choice was the modal one in 15 of the 18 pairs; black-black was the modal choice in no pairs. The median number of black-black plays was three whereas the median number of red-red plays was 28. If black and red choices were made at random with equal probability in the game, the red-red and black-black plays could be expected to vary around a mean of 12.5. In brief, none of the prison pairs displayed collaboration whereas 15 pairs could definitely be said to have employed a competitive strategy.

Another aspect of the prisoners' behavior was that red-red plays became more pronounced as the game progressed. The median number of red-red plays for the first 25 trials was 10; for the second 25 trials it was 16. The chi-square resulting from the medium test was 7.11, significant beyond the .01 level. (This progressive fixation on a competitive strategy was also true of the college group.) In brief, any attempts at collaboration in an effort to maximize joint return were quickly abandoned.

D. DISCUSSION

The theoretical interest in a game like the one we have employed is that there are arguments for playing either red or black. In a formal sense red dominates black in our game since no matter what the other player does, one still does better by playing red rather than black ($5 > 3$, $1 > 0$). For this reason Luce and Raiffa prescribe confession as the rational solution to the prisoner's dilemma even though confession by both produces a less desirable result for either participant than non-confession by both. However, the prisoner's dilemma in its originally stated form is a one-shot affair. Our version is an iterated game which affords the possibility of collusion between the players. According to Luce and Raiffa (1957 p. 101):

> If [the game] is played but once, we feel that it is "reasonable" to single out ($\alpha_2 \beta_2$) as the "solution" of the game provided that there is no preplay communication. . . . In contrast, we do not think it "reasonable" to single out [(α_2, β_2)] as the "solution" when [the game] is iterated n times. . . . It is not "reasonable" in the sense that we predict that most intelligent people would not play accordingly.[2]

As our data indicate, both college subjects and reformatory inmates play the game in an "unreasonable" way. For students the results may not be too difficult to understand. Supposedly, they live and work in a competitive atmosphere, and maximation of monetary return may very well have less utility than the desire to beat the other player, or at least, to avoid the possibility of being defeated by him. If one consistently plays red in the game, it is impossible to get less than the other player and if he plays black, one gets more, of course.[3]

---

[2] $\alpha_2\beta_2$ would correspond to red-red in our game.

With respect to the prisoners, there is a myth which holds that those who have been punished by society are motivated by a certain communality in interests. The present experimental situation is ideal for accumulating rewards at the expense of a bourgeois experimenter, but in no case did a pair employ a strategy that maximized joint return. It is true that the situation is one in which maximal joint return necessitates collaborative behavior from both since a black player has no choice but to defect to red, if the other player insists on playing red, but even initial, abortive attempts at playing black were notably absent.[4] In any event, a romantic or sentimental hypothesis to the effect that criminals tend to work out a mutually beneficial strategy when given the opportunity to exploit an outgroup member is not supported by these data.

The significantly greater number of red plays in the second half of the game has been found to be true of almost all our two-person non-zero-sum games. Our explanation is that in the early part of the game the money (or, in this instance, cigarettes) has more utility than it has as the game progresses. As the payoff decreases in utility, the competitive aspects of the situation become correspondingly strengthened.

E. Summary and Conclusions

Eighteen pairs of inmates at a Federal reformatory participated in a "prisoner's dilemma" game. No pairs displayed collaboration and 15 pairs employed a definitely competitive strategy. The behavior of the prisoners was almost identical with that of a college population in a similar game. The over-all results point to an unwillingness to engage in collaborative behavior in two very dissimilar populations.

---

[3] We have employed many versions of a prisoner's dilemma game with college students, and even some games in which black would dominate in the formal sense. We have yet to run an iterated two-person non-zero-sum game in which black-black is the modal play.

[4] "Cooperative" would be a more descriptive term than "collaborative," but it already has a special usage in game theory, namely, pre-play communication between players.

REFERENCES

1. DEUTSCH, M. Trust and suspicion. *Conflict Resolution,* 1958, **2**, 265-279.
2. LUCE, R. F., & RAIFFA, H. Games and Decisions. New York: Wiley, 1957.
3. SCODEL, A., MINAS, J. S., RATOOSH, P., & LIPETZ, M. Some descriptive aspects of two-person non-zero-sum games. *Conflict Resolutions,* 1959, **3**, 114-119.
4. VON NEUMANN, J., & MORGENSTERN, O. Theory of Games and Economic Behavior. (2nd ed.) Princeton Univ. Press, 1947.

Department of Psychology
Ohio State University
Columbus 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 139-144.

SUBLIMINAL TACHISTOSCOPIC PERCEPTION AS A FUNCTION OF THRESHOLD METHOD*

University of California, Los Angeles; and U.S. Veterans Administration Hospital, Los Angeles

ALLAN E. EDWARDS[1]

A. PROBLEM

Experiments relating behavior to some tachistoscopically presented variable have directed little attention to methodological considerations which are inherent to the technique used.

While *how much* the subject perceives has long been considered a function of method in psychophysics, e.g., an absolute threshold by the method of limits is not necessarily an absolute threshold by the method of constant stimuli, the consideration seems to have been overlooked in tachistoscopic work because of the acceptance that the subject is giving "verbal" or "recognition" responses. Bricker and Chapanis (1) have demonstrated the necessity for not accepting "recognition" at face value. Garner, Hake, and Erikson (2) point up the imperativeness of a closer look at "recognition," and Wilcott (4) provides a strong argument that the method of obtaining a limen must be considered before any stimulus can be called either sub or supraliminal.

If, concealed in the term "verbal response"/"recognition," lies a hierarchy of methods, rank ordered in terms of threshold points, then subliminal or supraliminal perception can be demonstrated at will by moving either up or down this ordered hierarchy. The demonstration of the relationships between at least some of these methods was the purpose of this experiment.

B. METHOD

1. *Apparatus and Subjects*

An episcotister tachistoscope, coupled through a minimax transmission to a constant speed electric motor, intersected the collimated light of a projector. The angular dimension of the open sector, and the velocity of the episcotister

---

* Accepted for publication by Roy M. Dorcus of the Editorial Board, received in the Editorial Office on May 10, 1960, and published immediately at Provicetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to express appreciation to S. Carolyn Fisher, University of California, Los Angeles, for constructive advice, and to Miss Margaret Gate for partially supporting this research.

could be varied such that one projection per second of controllable duration could be presented. A 300 watt, 35 mm slide projector, six feet from a large white bond paper screen with room illumination of approximately 1.6 to 1.8 foot lamberts was used throughout. The source of all stimulus words was the six-letter, category "A," (50-99 words per million) of the Thorndike-Lorge word count (3). The following words were used as stimuli: ARTIST, CASTLE, ENGINE, FLIGHT, GROWTH, INVITE, LOVELY, NOBODY, POLICE, POTATO, HARBOR, HUNGRY, SAILOR, SMOOTH, TEMPLE, WANDER, INSIST, COFFEE, MEMORY, and PARENT; printed in IBM executive type and negative projected. Subjects, 17 adults, ages 17 to 57, 4 females, 13 males.

2. *Procedure*

S's were seated comfortably in front of the screen. Once the experiment was underway they typically saw only a low frequency, low intensity flicker on the screen. A duration threshold was established by the method of constant stimuli to 50 per cent accuracy then altered by shortening the duration of each projection until that point where four words were consecutively presented 60 times each, *S* making at least 10 guesses to each word with no recognition apparent. From four to seven new words (from the word list cited) were then presented to *S*. The following tasks were demanded, in order, on the presentation of each word.

FV (Free Verbalizing). While visually fixating the flashing projection, *S* was required to verbalize in a free fashion by the usual (e.g., 6) instructions to elicit free associations. His verbal responses were recorded verbatim until he terminated of his own accord, producing approximately 80 words per stimulus.

DG (Discerning Guess). *S* made his three most discerning guesses of the word being presented. The instructions were to use any technique desired but to "go on record" with three guesses. The stimulus was present and conditions optimal for *S*.

MC (Multiple Choice). Projecting was then discontinued and *S* immediately chose a word from a handwritten list composed of the word projected and four randomly chosen foils from the same word source.

3. *Scoring*

The discerning guesses (*DG*) were evaluated as correct if the stimulus word appeared within any one of the three responses.

The multiple choice (*MC*) was evaluated as correct if the stimulus word was selected.

The free verbalizations (*FV*) were more difficult to evaluate, but were handled as follows.

After *all* data common to those words recognized by the *DG* method were removed, the *FV* data were treated in the following manners. The 67 remaining stimulus word—*FV* pairs—were presented to 10 judges, (3 psychiatrists, 3 psychologists, 2 vocational counselors, 2 graduate clinical trainees, all with U. S. Veterans Administration), with the following instructions:

> You are requested to make independent evaluations of the enclosed data. Take your time, look through all of it before you begin.
>
> Free-associative content was obtained from 17 subjects under impoverished stimulus conditions. Your task will be to decide how much, if at all, the said content (typewritten passages on cards) is related to its respective handwritten counterpart on the data sheet.
>
> Assume nothing, these may or may not be related. Judge by the following criteria.
>
> Mark one of the following categories in the appropriate place on the data sheet.
>
> *D*. No clear connection at all to word, as unrelated as any content would be. Very unlikely any relation between content and word.
>
> *CD*. In between.
>
> *C*. Some connection possible. Associated content might conceivably be related. Can't exclude altogether but relatedness seems unlikely.
>
> *BC*. In between.
>
> *B*. Connection probable. Associated content fits meaning of word very well. Seems meaningfully related and quite relevant to word.
>
> *AB*. In between.
>
> *A*. Close connection. Content wholly relevant to word. Fits word so well it could easily be just an extension of it.

The typewritten passages referred to in the judges' instructions were the *FV* responses of subjects and the handwritten counterpart were the stimulus words. The order of the stimulus word—*FV* pairs was randomly fixed for each judge and half of the pairs, randomly determined for each judge, were judged as obtained, *i.e.*, the *FV* datum for each word was paired with the actual word projected when the datum was obtained; the other half of the pairs was scrambled by a random procedure, i.e., each *FV* datum was paired with a stimulus word randomly sampled from the pool of presentations made to the *S*'s. The judgments were made independently and with no information as to the design of the judging task.

C. RESULTS

A total of 82 words were projected to the 17 *S*'s. It is worth repeating at this point that the projection duration for each *S* was where he evidenced no recognition, stimulus present, 10 guesses having been made to each of four successive words. No *S* perceived more than two words by any one evaluated method and no word was perceived more than expected from its proportion of presentation.

DG. As can be noted in Table 2, in 15 instances the stimulus word was recognized by this method, about what would be expected by altering the 50 per cent threshold. All *FV* and *MC* data common to these *DG* recognitions were then removed. By criterion of the operation used for establishing the duration threshold and by the criterion of *DG* recognition, the remaining worlds could clearly be classed as having been presented at subthreshold duration. They are treated below in the *FV* and *MC* evaluation.

FV. For statistical purposes the judgments placed in double lettered categories were divided equally into their adjacent categories yielding Table 1.

TABLE 1
FREQUENCY OF JUDGMENTS, 10 JUDGES, 67 STIMULUS WORD—FV PAIRS

Judgment category	Scrambled pairs	As obtained pairs
A	2.50	12.50
B	23.50	30.50
C	83.00	103.00
D	224.00	191.00

The chi-square of Table 1 is 12.16, $p < 01$. The *FV* data, as obtained from the *S*'s were more relevant to the actual projected words than would be expected with chance connections to the same words. All *MC* data common to instances where at least two judges had classed a word—*FV* pair as *BC* or above were removed and the remainder treated in the *MC* section below. By the criteria of the original operation, the *DG* and the *FV* operations this remainder should clearly be classed as subliminal.

MC. Table 2 notes that in 27 cases of these 57 the word was correctly selected by the *MC* method. The expected number of selections equals 11.40. Evaluating the obtained 27 against this number by the binominal, $p < .001$.

TABLE 2
FREQUENCY OF RECOGNITION BY METHOD; *N* CORRECTED AFTER EACH ROW; NUMBER OF WORDS RECOGNIZED

DG	FV	MC		
15	22	43	Total *N*	82
	10	34	Total $N - DG$ recognized	67
		27	Total $N - DG$ and *FV* recognized	57
			Unrecognized	30

D. INTERPRETATIONS

The rank ordering of the sensitivity of different kinds of "threshold" determinations makes so-called "subliminal perception" a methodological artifact. Establishing a "threshold" by an operation such as the *DG* method and later testing the *S* by either a variation of the *FV* or *MC* methods will yield subliminal perception but only because of differential sensitivity of these methods. A threshold gained by the *MC* method would yield supraliminal perception in either the *FV* or *DG* methods.

The *DG* method, the most common, is hardly as discerning as its name would indicate. *S*'s can discriminate a stimulus word more critically by at least using either of the other methods tested here. It is interesting to note that out of the 82 words presented, only approximately 30 of them were not recognized by at least one of the methods. The rank ordering of threshold *DG*, threshold *FV*, and threshold *MC* are particularly relevant to interpreting data asserting to establish subliminal perception.

The striking analogy between the results of this investigation and the well validated differences obtained by varying method in measuring Memory (5) is notable. Memory, as measured by recall, is quite different from the same as measured by recognition.

Operationally, the responses of a subject to a tachistoscopic exposure may as well be considered a function of short term memory as of perception. It is in this sense that the results are interpreted, there is no more reason to speak of "subliminal perception" than there is to concern ourselves with "subliminal memory." The quantity of memory and/or perception is a function of the measuring technique.

E. SUMMARY

Different methods of establishing thresholds to tachistoscopically presented words were tested on 17 *S*'s. The most sensitive was a multiple choice task, the least sensitive the discerning guess, "what-is-the-word?" task, and free-verbalizing was intermediate in sensitivity. It is suggested that "subliminal perception" is an artifact of establishing a threshold by one method then testing by another, more sensitive, method.

REFERENCES

1. BRICKER, P. D., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181-188.
2. GARNER, W. R., HAKE, H. W., & ERIKSON, C. W. Operationism and the concept of perception. *Psychol. Rev.*, 1956, **63**, 149-159.
3. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Columbia Univ., 1952.

4. WILCOTT, R. C. Subliminal stimulation vs. psychophysical thresholds. *Percept. & Motor Skills*, 1957, **7**, 29-36.
5. LUH, C. W. 1922, The conditions of retention. *Psychol. Monog.*, 1922, No. 142.
6. WOODWORTH, R. S., & SCHLOSBERG, H. Experimental Psychology. New York: Holt 1954.

Psychology Service
Wadsworth G.M. & S. Hospital
Veterans Administration
Los Angeles 25, California

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 145-152.

BINOCULAR RIVALRY OF PARENTAL FIGURES IN THE HUMASCOPE*

Department of Psychology, Dickinson College

H. M. WELLS AND D. M. BELL

A. INTRODUCTION

Lately a great deal of attention has been paid to the psychological aspects of perception. This has been particularly true in the area of "perceptual distortion" (Ames, 1952). It is suggested that perceptual distortion cannot be explained simply in terms of the standard stimulus properties presented to the organism. In recent studies of binocular rivalry and fusion (Engel, 1956; Bagby, 1957; Pettigrew, Allport, and Bernett, 1959; Hastorf and Myro, 1959; and Kohn, 1960), the meaningful content of the stimuli has been introduced as a major consideration. Perception is then described as a kind of directional activity in part determined by the past experience of the organism with the stimulus considered more as a probability than a rigidly defined substance with inherent properties.

Binocular perception with the use of stereoscopic techniques i.e., two separate stimuli simultaneously presented monocularly to each eye, are particularly adaptive for these kinds of considerations. In this experiment, the subjects were simultaneously presented with a slide picture of each parent to determine what kind of binocular fusion or rivalry might possibly occur as a function of the meaning each subject attached to the stimuli.

B. METHOD

1. *Apparatus*

The major piece of equipment used for this experiment was a machine designed by Adelbert Ames, Jr., called a humascope (Adlerstein, 1958). Ames and Hadely Cantril received two grants (Contract N onr-496(ol) with the Institute for Associated Research, Hanover, N. H., and Contract N60nr-27014 with Princeton University, Princeton, N. J. for research in perceptual distortion. This machine is an enlarged stereoscope consisting of a binocular viewing device with an adjustable knob for focusing, two separate chambers with a fluorescent light and two mirrors facing each other at ad-

---

* Received in the Editorial Office on May 11, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

justable angles for focus, a control for lighting with which the illumination in each chamber may reach equivalent or differential intensity, an adjusted timer which controls the time exposure as seen by the subject through the stereoscopic viewer, and two openings with chin rests which can be adjusted to points of equivalent heights for human beings to look through, or in which glass screens are placed for the projection of sides.

In this experiment slides were projected by two slide projectors onto ground glass screens. These slides were focused by the use of two pattern cards which had lines indicating the points of intersection of the bisectors of the angles of the card, so that the eyes of the figures projected coincided with the point of intersection. The subjects were seated on the reverse side of the humascope. The light was adjusted for equivalent intensity for each chamber and kept constant for each subject.

2. *Subjects*

A total of 49 subjects, 19 girls and 30 boys, was selected randomly from the freshman class. The ages of the subjects ranged from 18 to 20 years. At the beginning of the school term during registration week the parents of the subjects were asked to cooperate in an experiment by having their pictures taken. All the parents' pictures were taken on the same day under identical photographic conditions.

3. *Procedure*

Each subject was contacted individually and asked to cooperate in an experiment of perception. The subjects were brought into a room one at a time and asked to sit down, look into the humascope, and adjust the test pattern on his side of viewing by turning the focusing knob so that the vertical and horizontal lines intersected to form a cross. (The test pattern for the subject consists of one card, the right card, with a horizontal line in the center; and the one card, the left card, with a vertical line in the center.) The subjects were then given the following instructions:

"Please look into the machine again and tell me what you see."

The first 11 subjects were given a two-second exposure of the simultaneous presentation of the photographs of their mothers and fathers. The remaining 38 subjects were exposed to several time exposures ranging from one-fourth second to three seconds to determine discriminative variations at different thresholds.

All of the subjects' comments were tape recorded. The responses of the subjects were then categorized into four separate areas: (*a*) total blocking of one parent; (*b*) superimposition or fusion of a parent; (*c*) an alternation of one parent to the other; and (*d*) a dominance of one parent over the other parent also present. If it was not possible to determine into which of these categories the subject's response should be placed, he was asked the additional question:

"Please describe what you saw a little further."

In other words, when a subject saw consistently only one parent, and replied, "I saw my mother," or "I saw my father," his response was categorized as a total blocking (Category 1). When the subject clearly indicated that he saw a superimposition or fusion such as, "I saw both my parents represented equally in one picture," his response was categorized as fusion (Category 2). If the subject said that he saw both parents but at different times such as, "I saw my mother first, and then she faded out and I saw my father," and maintained a shifting pattern of response, his response was categorized as alternation (Category 3). When the subject indicated a response of a clear dominance of one parent such as, "I saw mostly my father but my mother's hair was there," the response was categorized as dominance (Category 4).

C. RESULTS

The following results were obtained from the subjects verbal reports. Table 1 indicates that nine of the 49 subjects consistently blocked out one

TABLE 1
NUMBER OF SUBJECTS WHO TOTALLY BLOCKED OUT ONE PARENT

S's	Sex	Time	Parent not seen
10	Female	2 seconds 1 second	Mother Mother
11	Female	2 seconds 1 second	Mother Mother
17	Female	1 second	Mother
19	Male	1 second	Mother
26	Male	1 second	Mother
36	Male	1 second 3 seconds	Mother Mother
38	Male	1 second	Mother
12	Female	1 second 3 seconds	Father Father
42	Male	1 second	Father
9	5 Male 4 Female	13 exposures	7 blocked Mother 2 blocked Father

parent, or 18.37 per cent of the subjects did not see one parent. Because of the consistency of report between one and two seconds, a one-second exposure was standardized as the first exposure presented to the subjects. It was discovered that variations do occur, however, between one second and shorter exposures, but that these changes did not occur with those subjects who blocked out one parent at the one-second threshold. Seven of the nine subjects who blocked out one parent did not see the photograph of the mother! The other two saw only their mothers and blocked out their fathers.

Table 2 shows that 18 of the 49 subjects saw a fusion or superimposition of both parents in one picture, or 36.73 per cent of the subjects gave a "normal" response to the slide presentations.

TABLE 2

NUMBER OF SUBJECTS WHO SAW A SUPERIMPOSITION OR FUSION

S's	Sex	Time exposure
2	Female	2 seconds
3	Female	2 seconds
4	Female	2 seconds
5	Female	2 seconds
7	Female	2 seconds
13	Female	1 second
14	Female	1 second
23	Male	1 second
25	Female	1 second
30	Male	1 second
32	Male	1 second
33	Male	1 second
41	Male	1 second
43	Male	1 second
44	Male	1 second
46	Male	1 second
47	Male	1 second
48	Male	1 second
18	8 Female 10 Male	18 Exposures

Table 3 indicates the number of subjects who saw first one parent and then the other, although they saw both parents in one exposure. Eleven of the 49 subjects fit into this category, or 22.45 per cent of the subjects were classified into the alternation category. Except for two subjects (15 and 27), the parent seen first was also the figure which dominated the photograph. In four cases (8, 9, 29, 40) the subjects saw one parent exclusively at time exposures less than one second when they had originally seen an alternating pattern of each parent. In each of these cases, the parent who appeared

TABLE 3
NUMBER OF SUBJECTS WHO SAW AN ALTERNATION

S's	Sex	Time exposure	Parent seen first	Indicated dominance
6	Female	2 seconds	Father	Father
		1 second	Father	Father
8	*Female*	2 seconds	Father	Father
		1 second	Father	Father
		¾ second	Father	Father
		½ second	Father only	
		¼ second	Father only	
9	Female	2 seconds	Father	Mother
		1 second	Mother	Mother
		¼ second	Mother only	
		¼ second	Mother only	
15	Female	1 second	Father	Mother
		½ second	Both	Father
20	Male	2 seconds	Mother	Mother
		1 second	Mother	Mother
		1 second	Mother	Mother
		½ second	Mother	Mother
21	Male	1 second	Mother	Mother
		½ second	Mother	Mother
22	Male	1 second	Mother	Mother
27	Male	1 second	Father	Mother
		3 seconds	Both	Mother
		¼ second	Mother	Mother
28	Male	1 second	Mother	Mother
29	Male	1 second	Mother	Mother
		3 seconds	Mother	Mother
		¼ second	Mother only	
40	Male	1 second	Mother	Mother
		2 seconds	Mother	Mother
		¼ second	Father only	
		1 second	Mother	Mother
11	7 Male 4 Females	31 Exposures	15 Mother 8 Father 2 Both	19 Mother 6 Father

dominant at a longer exposure now appeared exclusively with the exception of Subject 40.

Table 4 indicates the number of subjects who saw both parents in one picture with a definite predominance of one of the two. Eleven subjects, or 22.45 per cent of them, saw both parents with one in dominance.

The data indicate that binocular perception of parental figures presented by sides in the humascope varies with the individual. Using a stimulus-response hypothesis all 49 subjects should have seen both parents in one

TABLE 4

NUMBER OF SUBJECTS WHO SAW BOTH PARENTS BUT WITH ONE PARENT PREDOMINATING

S's	Sex	Time exposure	Dominant figure indicated
1	Female	2 seconds	Father
16	Female	1 second	Father
18	Female	1 second	Father
24	Male	1 second	Father
31	Male	1 second	Mother
34	Male	1 second	Mother
35	Male	1 second ¼ sec. 3 seconds	Mother Mother Mother
37	Male	1 second 3 seconds ¼ sec.	Father Father Father
39	Male	1 second	Father
45	Male	1 second	Mother
49	Male	1 second	Father
11	8 Male 3 Female	15 Exposures	9 Father 6 Mother

picture. This is not taking into consideration possible lack of correction for eye dominance. Equal intensity of light and the presence of the test pattern before the experiment (indicating that all subjects perceived a portion of each of the two presentations) were the only corrections used. Nonetheless, the fact that 18.37 per cent of the subjects did not see any portion of the one stimulus which undoubtedly had some significance for the subject further substantiates the theory of psychological selectivity. This result also gives further support to other experimental work which demonstrates that subjects tend to structure their perceptual responses by perceiving that which is of dominant value to them.

Further research and experimentation should be carried out using a shorter threshold, one-half and one-fourth second exposures, to determine the relevance of a first impression or a strictly perceptual response, before the cognitive processes play a major role and the subject makes a judgment. As is noted in the tables, although perceptual differences did not occur between those subjects who reported a dominance of one parent at one second and their report at shorter exposure times, variations in their perceptual responses did occur for those subjects who perceived an alternation at a one-second exposure and at shorter exposures.

The data indicate that 63.27 per cent of the subjects involved in the experimentation did not have perceptual responses corresponding with the formal stimulus properties of the stimulus presentation.

D. Discussion

Several experiments in binocular perception on rivalry and fusion have indicated that meaningful content plays an important role in the physiological perception of a stimulus, (Hastorf, Engel, Bagby.) This experiment further confirms these findings. The prior experiments were conducted with a small stereoscopic device developed by Engel. This experiment is one of the few first attempts to use the humascope (an enlarged stereoscope developed by Adelbert Ames, Jr., and Hadley Cantril) for any systematized study.

The lack of correspondence between a subject's perception of a stimulus and what appears to be stimulus properties cannot be dismissed at a "perceptual illusion." The only interpretation which appears feasible for this phenomenon requires the introduction of dominance or meaningful content associated with this stimulus by the subject; i.e., the subject in part creates his own stimulus. This theoretical approach has been proposed by Hadley Cantril and has been called a transactional approach. This point of view discards the idea that perception can be explained in terms of a fixed stimulus and a response to that stimulus. The transactionalist emphasizes the importance of past experiences which the subject brings to the present situation. It is believed the meaning or purpose the subject has attached to this particular experience will affect his actual perceptual recognition of the stimulus, a kind of weighted average of past experience relative to the individual (Cantril). In other words, in the case of binocular perception, the individual will exclusively see or dominantly see that which has the most meaning for him at that time.

There is strong evidence from the field of social psychology that an individual tends to structure an ambiguous situation. The way in which the individual structures the situation is meaningful and is probably related to the significances based on past experiences which the individual externalizes. By structuring an ambiguous situation, the individual seems to give himself external reality which he can grasp and give meaning to.

One cannot adequately determine in this experiment why the one parent appeared to be more significant to the individual at the time. Further measures of other variables should be used to test the hypothesis of perceptual defense. For example, perceptual blocking may have a definite relationship with authority figures. It may also have diagnostic value for psychiatric institutions or clinics in which the basic emotional problems of the individual are related primarily to his relationship with the parental figures.

E. SUMMARY AND CONCLUSION

An experiment of binocular perception dealing with the rivalry and fusion of parents was conducted with 49 freshman students using the humascope. After having been shown time exposures of a slide presentation of both parents presented simultaneously, 36.73 per cent of the subjects saw a superimposition of both parents or a fusion, 22.45 per cent of the subjects saw both parents with a definite dominance of one parent, 22.45 per cent of the subjects saw both parents in one exposure but with alternation of the particular parent seen, and 18.37 per cent of the subjects totally blocked perceptual recognition of one parent entirely during a one- or two-second exposure. The evidence seems to indicate that perceptual selectivity is present in the simultaneous presentation of two stimuli with meaning, and that the one of dominant value is that which is most readily perceived.

REFERENCES

1. ADLERSTEIN, A. M. The humascope: a modified stereoscope. *J. of Psychol.*, 1958, **45**, 109-113.
2. BARTLEY, S. H. Principles of Perception. New York: Harper 1958. (Pp. 22-25.)
3. CANTRIL, H., & BUMSTEAD, C. Reflections on the Human Venture. New York: New York Univ. Press, 1960. Pp. 58.
4. ENGEL, E. The role of content in binocular resolution. *Amer. J. Psychol.*, 1956, **69**, 87-91.
5. ITTELSON, W. H. The Ames Demonstrations in Perception. Princeton: Princeton Univ. Press, 1952. Pp. 50.
6. KILPATRICK, F. P. Recent Transactional Perceptual Research—A Summary: Final Report on Office of Naval Research Contract N6onr-27014 with Princeton University Press. Princeton, N. J., 1955.
7. KOHN, H. Some personality variables associated with binocular rivalry. *Psychol. Rec.*, 1960, **10**, 1-7.

Department of Psychology
Elmira College
Elmira, New York

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 153-171.

THE MULTIFACTOR-ANALYTIC THEORY OF EMOTION*

Department of Psychology, Hofstra College

ROBERT PLUTCHIK[1]

A. INTRODUCTION

A theory of emotion must somehow integrate in a meaningful way the known observations in the field and at least suggest the path along which answers may be obtained to unresolved questions. In the study of emotion, there have been many problems raised which have received only scant or partial answers, such problems for example, as: What is an emotion? Is emotion the cause of a bodily state, the bodily state itself, or the result of a body state? Can there be emotions without awareness? To what extent do maturation and learning affect emotion? Are emotions disruptive or organizing?

Nearly every author who has attempted an answer to any of these questions has noted the frequent and persisting disagreements in the field. For example, Woodworth and Schlosberg (1954) write: "The topic of emotion has perhaps generated more unprofitable controversy among psychologists than any other with which they have concerned themselves." Hebb (1949) adds: "The discussion of emotion has been about as confused as that of any topic in psychology, partly because the terminology is often equivocal and partly because tradition carries great weight in this part of the field." In his book on *Emotions and Memory*, Rapaport (1950, p. 11) indicates that: "The main difficulty in the literature of emotions appears to be that the word "emotion" is sometimes used to designate a *phenomenon*, and sometimes to designate the *dynamics* underlying a phenomenon or group of phenomena." At last, but certainly not least, Leeper (1948) writes that, "in our thinking about emotion, . . . we have been dominated by *a priori* conceptions. Too commonly we have adopted, for all practical purposes, a faulty representation of psychological functions which was started back in the late 1700's—a division of psychological processes into those of cognition, affection, and conation." These four examples were cited mainly to emphasize the sense of confusion which

---

* Received in the Editorial Office on May 19, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is an expanded version of one delivered as part of a symposium on, "Current Status of Theories of Emotion," at the American Psychological Association meeting in Cincinnati, Ohio, September, 1959.

seems to permeate the field and to stress the need for a reorganization of thinking about emotion.

Some of the reasons for these confusions are related to the differences of approach of the clinician as contrasted with the experimental psychologist, some to the difficulties of experimentation in this field, while still others are based on a faulty analysis of what a definition of "emotion" requires (Plutchik, 1955, 1957). It is also evident that a good deal of research on emotion has been too narrowly focused on only one or two emotions, such as fear and anger, out of the total gamut possible. In this connection, Leeper (1948) writes, "unless we take an extremely pessimistic view of human life, we might well say that such 'pleasurable' emotions or 'positive emotions' are, in general, just as numerous and important in human life as are the 'unpleasurable' or 'negative' emotions. It is hard to see, therefore, why they secure mainly passive mention and why such emotions as fear and anger are discussed as though they are the only valid prototypes of emotion." Any adequate theory must include the entire range of emotion in the same framework. In the next section a brief review of six theories of emotion will be made as an introduction to the view to be presented here.

B. Review of Other Theories

1. *The James Theory*

The classic theory in the field is the formulation of William James, who stated his view in the following terms, ". . . bodily changes follow directly the *perception* of the exciting fact, and our feelings of the same changes as they occur *is* the emotion" (James, 1884). It is clear from his writings that he conceived of his theory largely to answer the one question, 'What comes first in emotion, the feeling or the bodily changes?' His answer has been challenged and defended by many psychologists since that time but a definite answer to the question is still lacking. In fact, Lindsley, in his article on Emotion in the Handbook of Experimental Psychology, writes: "It appears that the James-Lange theory, by its very nature and formulation, is not a theory but an untestable hypothesis" (Lindsley, 1951). The theory does not even attempt to answer most of the questions listed earlier. In addition, the formulation does not deal at all with the problem of acute versus persistent emotions, nor with the question of repressed emotions, i.e., emotions without awareness. This last possibility would be meaningless in James' terms and yet emotions without awareness are a commonplace of clinical practice. It is of some interest to note in this connection that Freud made one comment on the James-Lange theory in his published writings, to wit, "what psychology has to say about affects—the James-Lange theory, for instance—is utterly incomprehensible to us psychoanalysts and impossible for us to discuss" (Freud, 1953, p. 404).

2. *The Attitude Theory*

Within the last few years, a theory of emotion has been proposed by Nina Bull which has certain similarities to the James-Lange hypothesis—the so-called Attitude theory of emotion. In the author's words: "The fundamental concept in the theory is that feeling or affect is dependent on preparatory motor attitude— and hence that attitude is essential to emotion. . . . Emotion is conceived of as a sequence of neuro-muscular events in which a postural set or preparatory motor attitude is the initial step. This preparatory attitude is both involuntary and instinctive and is the end-result of a slight, tentative movement which gives a new orientation to the individual, but does not immediately go into the consummatory stage of action" (Bull, 1952). In another paper she adds that the difference between this and the James-Lange theory, for example, is that "the sorry feeling comes with readiness to cry, and not because of actual crying. . . . Violent crying eliminates the sorry feeling" (Bull,1945).

In two ingenious studies, Bull and her co-workers, using hypnotized subjects, were able to find certain more or less consistent bodily sensations reported when certain emotion words such as fear, anger, disgust, triumph, or depression were suggested (Bull and Gidro-Frank, 1950). Later, when other subjects were asked to "lock" their postures into certain positions, again under hypnosis, they could not change feeling unless their positions were also changed (Pasquarelli and Bull, 1951). Two important conclusions stemmed from these findings: first, that skeletal muscle activity as well as visceral activity is involved in emotional feeling, and secondly, that different emotions not only feel different but have different postural sets associated with them.

Although the theory and experiments represent an interesting development of the James approach, a number of questions are left unanswered. Those concerned with the role of learning and with the differences between acute and persistent emotions as well as the question of emotions without awareness are particularly challenging here.

3. *The Cannon-Bard Theory*

A view which has achieved the greatest prominence in the analysis of emotions is the Cannon-Bard theory. In the words of the author, "when (thalamic) neurones discharge in a particular combination they not only innervate muscles and viscera but also excite afferent paths to the cortex. . . . The peculiar quality of the emotion is added to simple sensation when the thalamic processes are roused" (Cannon, 1927). The evidence for this belief has been obtained mainly through extirpation studies. Yet, more than 20 years after the initial statement of the theory, Lindsley could write in the *Handbook* (Lindsley, 1951): "With increasing knowledge of the central and peripheral nervous systems and their functions, it becomes apparent that previous theories of emotion (James-Lange) or of emotional behavior (thalamic theory) are inadequate to account for not only the peripheral bodily changes in emotion, but the central manifestations as revealed by electrocortical activity." When we add to this the many criticisms that have been made of the thalamic theory by Masserman (1950), Magda Arnold (1950), and Hebb (1949) to name only a few, we can hardly accept this theory as a demonstrated truth. From another point of view, the thalamic theory, even if true, could at best provide us only with a statement of certain neural structures which are involved in emotional expression; this is hardly an advance over saying that the autonomic nervous system participates in emotional expression. Almost all of the basic questions that were originally listed are left untouched by the thalamic theory. As Rapaport points out, ". . . . the idea of one center for such a complicated function as emotion is alien to the present conception . . . that functions are represented on different neural levels and depend on an axis of centers rather than on one center . . . In the present state of knowledge the problems of emotion cannot be solved by a physiological approach only" (Rapaport, 1950). Freud's comment is especially relevant here too: "Today I must say I know of nothing less important for the psychological comprehension of anxiety than a knowledge of the nerve-paths by which the excitations travel" (Freud, 1953, p. 401).

4. *The Activation Theory*

A fourth theory which has been proposed is called the Activation theory of emotion and has been mainly developed by Harold Schlosberg (1954): "Instead of treating emotion as a special state, differing qualitatively from other states, the theory locates emotional behavior on a continuum that includes all behavior. This continuum, general level of activation, has its low end in sleep, its middle ranges in alert attention, and its high end in the strong emotions." In addition to this intensity dimension two qualitative dimensions are added; pleasantness—unpleasantness, and attention—rejection. These three dimensions of emotion are conceptualized by a cone-shaped model with a roughly circular surface depicting the two qualitative axes, while the vertical dimension represents the intensity dimension. Thus far, the only experimental work based on the model concerns the judgment of facial expressions from photographs using the two qualitative axes, and the effect of level of activation on certain simple motor tasks and GSR.

This theory is an important advance over the previous ones in that it does attempt to integrate into a single model several recognized dimensions of emotion; it also includes pleasant as well as unpleasant feelings. It does not, however, explicitly comment on a number of important issues which have been raised here. For example: what is the difference between acute and chronic emotion in this theory; what is meant by repression of emotion; what produces emotion; what *is* an emotion; and so on. It is quite possible, of course, that the answers to these questions can eventually be made explicit by the theory—this in fact could be true of any of the theories already described—but these questions have not received an explicit answer as yet.

Lindsley (1951, p. 509), the author of the term "activation theory," concludes his *Handbook* article with the following statement: "The activation theory appears to account for the extremes, but leaves intermediate and mixed states relatively unexplained as yet."

5. *The Psychoanalytic Theory of Emotion*

The fifth theory to be reviewed is that proposed by David Rapaport (1950) in his book *Emotions and Memory;* he calls it a psychoanalytic theory of emotion. Rapaport quotes Freud (1925) as saying that "affects and emotions correspond with processes of discharge, the final expression of which is perceived as feelings. In the present state of our knowledge of affects and emotions we cannot express this difference more clearly." Those investigators who seek to establish clearcut relations between "emotion felt" and bodily processes are bound to fail because "the dynamics of psychic manifestations are *unconscious,* and cannot be found by investigating interrelations of the data of physiology and the data of consciousness" (Rapaport, 1950, p. 29).

The old source of controversy between the James-Lange and the Cannon-Bard views is essentially dismissed by pointing out that in the psychoanalytic theory, a percept will elicit unconscious processes which may be followed in any sequence by the bodily process or by the "emotion felt" or by only one of these or by neither; for both are conceived of as manifestations of the same psychic process. Emotion as well as thinking according to Freud's theorizing (1925) is experimenting at action with small amounts of energy.

The following main elements are *assumed* in the psychoanalytic view: (*a*) that an unconscious process occurs between the perception of the stimulus evoking emotion and the peripheral physiological process: (*b*) that the peripheral physiological process and the "emotion felt" are both discharge processes of the same instinctual source of energy; (*c*) that emotions are expressions of instinctual conflict; and (*d*) that momentary manifestations of emotion should be differentiated from chronic.

This theory, as far as it goes, makes some important contributions. It eliminates the sequence problem (i.e., which came first, the feeling or the bodily change), an issue which has proved an unanswerable stumbling block for over 50 years; it emphasizes that emotional processes are unconscious and they cannot, therefore, be examined by introspection alone—by implication these processes must be inferred in human beings, just as we infer emotional processes in lower animals. It suggests that conflicts are involved in emotion and recognizes the problem of persistence.

The main weaknesses are its vagueness and metaphorical character, and the fact that it leaves unanswered certain basic questions. For example, what does it mean to control our emotions, how do the emotions of infants differ from those of adults, what are the stimuli which produce emotion, are there differences between emotions and so on. There is also the possibility that different analysts have somewhat different conceptions about the nature of emotions; there is no unanimity as yet, nor was there in Freud's own writings. There is thus adequate opportunity for development of these conceptions.

6. *The Motivational Theory of Emotion*

A few years ago Leeper (1948) published an article in which he advocated a motivational theory of emotion to replace the definition of emotion as a disorganized response. He pointed out that the most common definition of emotion is exemplified by Young's description (1943, p. 60), i.e., "an acute disturbance of the individual as a whole, psychological in origin, involving behavior, conscious experience, and visceral functioning." Leeper shows a number of contradictions in the previous uses of the term disorganization and finally suggests a specific definition, to wit: "The criterion of organization, consequently, is not a matter of whether there is some interference with preceding activities. It is the question whether this interference is relatively chaotic and haphazard, or whether the suppressions and changes of subordinate activities are harmonious with some main function which is being served."

Based on this definition emotions can be conceived of primarily as motives because they are processes which arouse, sustain, and direct activity. The physiological and behavioral changes that occur during emotions may be different from previous states, but they help the individual prepare for, or do what is necessary at the moment.

This conception is in certain respects consistent with the idea that emotions represent "emergency" reactions of fight or flight, in that the preparation for such energetic activity is highly organized from a physiological point of view, even though preceding activities such as digestion or intellectual activity, may be modified or temporarily stopped. Leeper's main point of disagreement with the emergency theory is that he insists that pleasurable emotions such as joy, love, and delight do not represent emergency reactions, but that they still must be included within the framework of a theory.

This emphasis on emotion as organized is a useful antidote to earlier rationalistic speculations which conceived of emotion only as interfering with activities. However, as Hebb (1949, p. 240) points out, the result of Leeper's argument is to broaden the category of emotion so much that it includes all psychological processes because almost all act to "arouse, sustain, and direct" behavior. Hebb suggests as an alternative view that emotion be regarded not as a single, unitary process, but in dualistic terms: *integrative* emotions are those in which there is a tendency to maintain or increase the original stimulating conditions, while *disintegrative* emotions are those in which the tendency is to abolish or decrease the stimulus.

Although this problem of organization in emotions is an important one, it is still too narrow to qualify as an adequate theory of emotions. Many other important problems concerned with the effects of learning, the nature of patterning, the nature of repression and control are left unanswered. An important issue has been touched but a comprehensive theory is still needed.

7. *Summary*

The preceding pages have briefly outlined six different theories of emotion. Each has tried to deal with some important problem and yet at the same time has left unanswered other important problems. Some of these theories have persisted in our textbooks even in the face of contradictory facts or relevant criticism. Why this is so is probably due to the nature of the scientific process, for as Conant (1947) has remarked: "A theory is never overthrown by contradictory facts. It is only overthrown by a better theory." All the theories can be shown to contribute something of value in the attempt to develop a new integration.

C. The Multifactor-Analytic Theory of Emotion

In trying to evaluate the adequacy of any theory, important criteria to be considered are that known observations be accounted for, new observations suggested, and apparently separate ones are shown to be related. Another aspect of theory construction is that the basic terms or ideas of a theory need not necessarily accord with common sense, as long as the experiences of common sense can ultimately be derived from them. Neither electrons nor the libido of psychoanalytic theory are observed directly, but can only be inferred, just as pure colors and pure elements can only be approximated more or less closely. It should be noted also that most theories tend to use analogy as a guide for their development, or in other words, that a model is used as a basis for suggesting relationships that might not otherwise be obvious.

The theory of emotion presented here utilizes all these procedures. It attempts both an analysis and a synthesis of all the complex emotions of everyday life; it conceives of the units of analysis as hypothetical constructs, and it uses the analogy of the color solid to guide its development.

1. *Postulates of the Theory*

Postulate 1. There is a small number of "pure" or "primary" emotions.

Postulate 2. All other emotions are mixed, that is, they can be synthesized by various combinations of the primary emotions.

Postulate 3. Primary emotions differ from each other both with regard to physiology and behavior. There is a growing body of literature that indicates that there are definite differences between such emotions as fear and anger, disgust and depression, and pleasure and resentment. Such differences have been reported by Ax (1953), Davis and his co-workers (1955), Funkenstein (1955), Wolff (1950), Bull and Gidro-Frank (1950), and Pasquarelli (1951), to name a few.

Postulate 4. Primary emotions in their pure form are limiting states whose properties can only be inferred from various kinds of evidence. This is based on the fact that most, if not all, the emotions of daily life are mixed, and that many problems of measurement are involved.

Postulate 5. Primary emotions can be conceptualized in terms of pairs of polar opposites. Many emotion terms represent opposites, and factor analysis of temperamental characteristics usually produces bipolar factors.

Postulate 6. Each emotion can exist in varying degrees of intensity, or levels of arousal. Duffy (1957), Schlosberg (1954), and Malmo (1957) have provided experimental evidence for a level of activation dimension. Our language distinguishes between grief and sadness, terror and apprehension, and ecstasy and pleasure.

The color-cone analogy was used to organize the postulates. This choice of a model meant that the primary emotions should be arranged around an "emotion circle," that there should be a gradation of emotions (in some reasonable sense) around the circle, and that the primaries be at least subjectively "pure," that is, unanalyzable on an introspective basis.

After considerable examination, four pairs of primary emotions were decided on, each pair conceived as bipolar factors, and tentatively called by the following names: joy versus sadness, acceptance versus disgust, surprise versus expectancy, and fear versus anger.

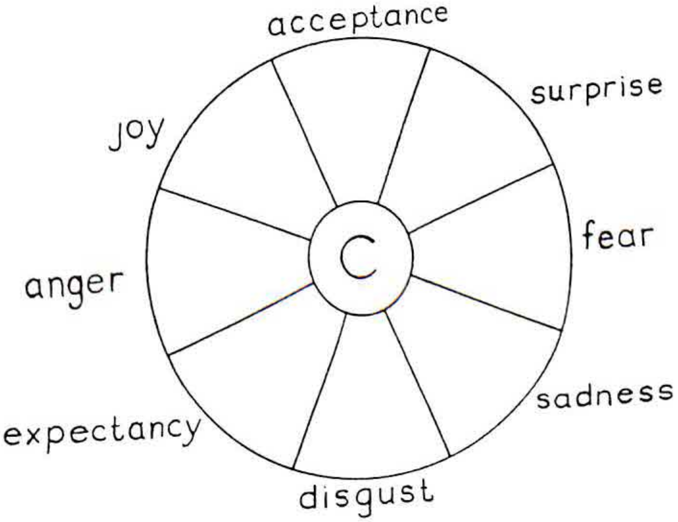

FIGURE 1
A Cross-Section of the Emotion-Solid

These are arranged around a circle as shown in Figure 1. Opposite emotions on the circle are complementary, as are colors; that is, we cannot conceive of opposite emotions as existing simultaneously without conflict, or inhibition. The letter "*C*" at the center of the circle is designed to represent this idea. It may also be noted that there is a certain reasonableness in the progression of primary emotions around the circle. Although joy is a unique emotion, it is expansive and excited as is anger, while both sorrow and fear are contracted. Surprise is related to startle, which is close to fear. Expectancy, set, or vigilance is an outwardly directed process, and thus belongs more closely to anger than fear or sorrow. Yet, just as each primary color is unique, so is each primary emotion unique.[2]

It would be of some value at this point, to justify in at least an heuristic way, the particular ordering of the primaries around the solid. Relevant evidence may be found in studies of facial expression and in studies of individual emotions as well as in the internal consistency of the theory.

In connection with the first point Woodworth and Schlosberg (1954) describe a circular scale for facial expression of emotion based upon a "confusion" technique. For example, judges tend to confuse the expression of "anger" with "determination," "disgust" with "contempt," and "love" with "happiness" and "mirth." This is an observation based on posed facial expressions (Feleky and Frois-Wittmann pictures). Based on this approach Schlosberg set up a six-category scale in the following order: love-happiness-mirth is one category; surprise is a second; fear-suffering, a third; anger-determination, a fourth; disgust a fifth; and contempt, a sixth.

There are several points that might be raised with regard to this data: (*a*) the naming of the categories is largely dependent on the arbitrary names of the expressions to be posed, decided in advance; (*b*) the actual grouping is partly arbitrary, for example, in Feleky's original data, fear and anger are confused more often than fear and suffering and yet the latter two are grouped together (Woodworth, 1938); (*c*) only the Frois-Wittmann pictures give evidence of a circular rather than a linear scale, since no circular scale is found in the original Feleky pictures nor in "candid-camera shots" taken from magazines (Woodworth and Schlosberg, 1954, p. 126).

It may be concluded that this kind of study of facial expressions suffers from problems common to most studies using posed expressions; these include (*a*) the number of categories of emotions used, (*b*) whether judgments should be free or restricted, (*c*) how to deal with the large vocabulary of emotion words obtained as responses, as well as (*d*) the criterion of a "correct" judgment.

Most of these problems were minimized or avoided in a study done by Hill (1955). A list was made of the various expressions the forehead, eyes, and mouth could assume, and then two subjects, a male and a female, were asked to set their face into these partial expressions one by one while a color slide was taken. The slides were then masked with black tape so that only the desired feature was visible and these were then shown to a group of adults who were to judge the emotion best represented by the various expressions on each part of the face, using only the eight primary emotion terms of the present theory. From this data it was possible to construct a composite picture best describing each primary emotion and to determine the degree of overlap of judgment. It was found that surprise was most often grouped with fear, disgust most often grouped with anger, and joy most often grouped with acceptance, results which are fairly consistent with the ordering around the emotion circle.

---

[2] Figures 1, 2, and 3 are used with permission of the Transactions of the New York Academy of Sciences.

One very important point should be made clear in this connection: the assumed primaries are hypothetical constructs, states of the organism as a whole, and are therefore only partially reflected by differences of facial expression. Only by a series of successive approximations can we build up a composite picture of these primaries as a totality.

Studies of the pattern of individual emotions are relatively few. In the hypnotic experiments described by Bull (1951), the pattern of fear which emerged was "a strong compulsion to escape, together with a general tensing up or freezing of the body," while anger was described as "a primary compulsion toward aggression or attack." This is, of course, consistent with Cannon's conception of emotions as emergency reactions with fear being characterized by "flight" and anger by "fight" (Cannon, 1939). Wenger, Jones, and Jones (1956) point out that fear is associated with sympathetic dominance and have "cold feet," cold hands, and paleness as typical concomitants, while anger is a "warm" emotion showing strong parasympathetic components. These observations suggest the opposite nature of fear and anger in terms of patterns of response. The other bipolar emotions seem to require less justification since they appear more intuitively opposite; this perhaps accounts in part for the fact that such little research has been done on them.

The third general justification for the ordering of emotions in the particular way shown in Figure 1, might be thought of in terms of the implications and internal consistency which results. Some of the internal consistencies are illustrated in Figure 3 by the fact that pride is opposite shame, aggression is opposite alarm, and love is opposite misery (not hate), while expectancy mixed with joy may be defined as "optimism," but when mixed with its opposite, sadness, produces "pessimism." Similarly, a mixture of acceptance and fear may be thought of as "submission," while acceptance and anger produces "dominance." Some of the other implications will be described in the following pages.

One problem which arises, is that of how to best name each primary emotion since, if the intensity or level of arousal dimension is considered, each primary emotion can exist in varying degrees. It is difficult to specify which particular term for one emotion (for example, fear) should be set approximately equal in intensity to terms for other emotions. The one set of exceptions would be the terms describing the primary emotions when activated to their highest degree. We should then have esctasy instead of joy, terror instead of fear, revulsion instead of disgust, and rage instead of anger. We may thus conceive of the model in three-dimensional terms roughly as half an orange with eight sections, as shown in Figure 2. If we imagine

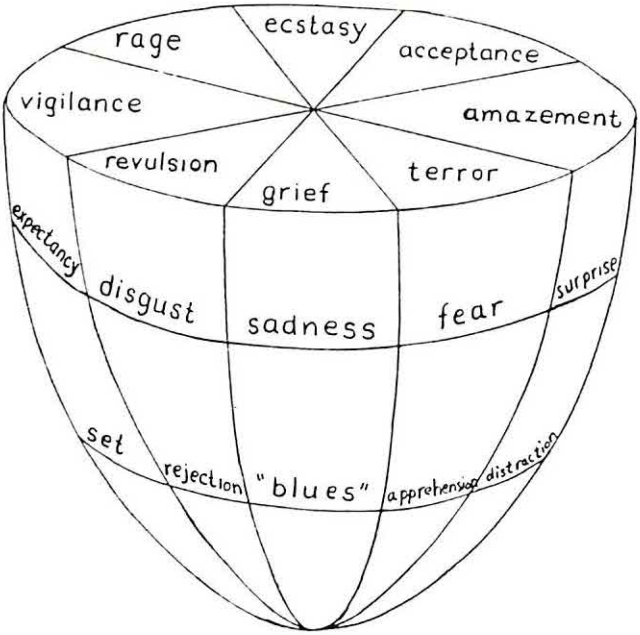

FIGURE 2
A Multi-Dimensional Model of the Emotions

taking successive cross sections, we keep duplicating the emotion circle with progressively milder versions of each of the primaries. The relative differences between them also decrease as the intensity level of sleep is approached. What the exact form of the emotion-solid is, must be an empirical question, and can be determined only by studies of level of arousal of the different primaries.

If the analogy with the color wheel is continued, it is possible to conceive complex mixtures of emotions synthesized by combining the primaries in various ways. A mixture of any two primaries may be called a dyad, of any three primaries, a triad. However, these dyads and triads may be formed in different ways. From Figure 1 it is evident that it would be possible to combine any adjacent pair of primaries to form an intermediate emotion, just as any two adjacent colors form an intermediate hue. Such adjacent mixtures of emotions are called primary dyads. Mixtures of two primary emotions which are once removed on the circle are called secondary dyads, and mixtures of two primaries that are twice removed on the circle are called tertiary dyads. The same general method of designation would apply to triads as well.

The problem of naming the mixtures that result is a difficult one. Because of the varying intensity levels there will be a certain element of arbitrariness. One criterion that is useful, however, is the requirement that opposite dyads tend to have opposing characteristics. This would be a criterion of internal consistency. Another approach might be to ask a group of judges to interpret the mixtures and to use the modal responses.

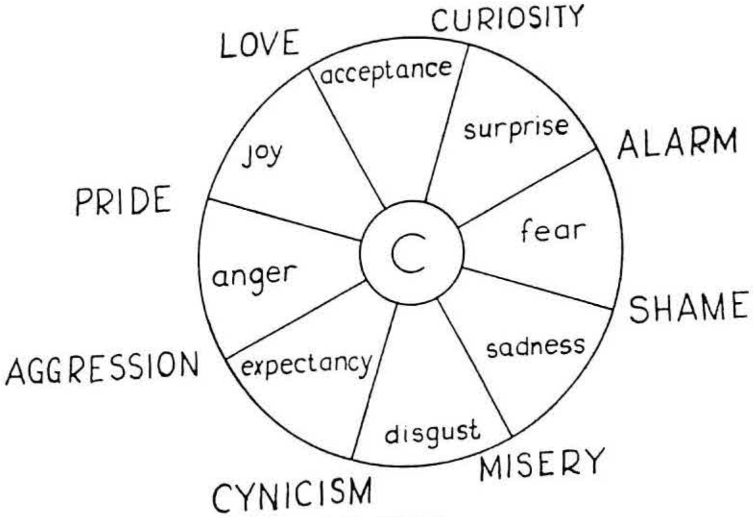

FIGURE 3
A CROSS-SECTION OF THE EMOTION-SOLID SHOWING THE PRIMARY DYADS

The list of primary dyads shown in Figure 3 is based on both these methods. The names attributed to these various mixtures are, of course, tentative, and it is not always easy to name each combination of emotions. This may be due to one or more of several reasons: perhaps our language does not contain emotion words for certain combinations, although other languages might; perhaps certain combinations do not occur at all in human experience, just as chemical compounds can be formed only in certain limited ways; perhaps we must make up our own names to describe certain compounds; perhaps we have not yet discovered all the emotional combinations of which man is capable. In chemistry, when a gap was discovered in the periodic table it was still possible to discuss the properties of this hypothetical substance, to name it, and eventually to discover it.

One possible way of dealing with the problem of naming emotion mixtures is to define them in terms of a "standard observer," as has been done for colors.

TABLE 1

This presents a tentative listing of the primary, secondary, and tertiary dyads. The terms in parentheses are alternative designations of the emotion mixture. The dashed line indicates that no suitable name has been found.

Primary dyads	Secondary Dyads	Tertiary dyads
Pride	Dominance	Resentment (Outrage)
Love	Delight	Guilt
Curiosity	Submission	Resignation
Alarm	Disappointment	Disgrace
Shame	Horror	Anxiety (Dread)
Misery	Scorn (Contempt)	Sullenness (Envy)
Cynicism	Optimism	— — — —
Aggression	Pessimism	Fatalism

In Table 1 is listed, tentatively, all the dyads: primary, secondary, and tertiary. An examination of this table reveals several interesting points. First, almost all of the dyads listed represent emotions that are considered to be relatively persistent. Emotions such as pride, aggression, submission, and optimism are usually long-lasting and, in fact, are often called personality traits. We talk of proud, aggressive, submissive, and optimistic *people*. This suggests that the formation of personality traits is related to the development of mixed emotions, and that persistent emotions are by their very nature as mixtures, quite different from the primary ones. This implication of the theory is completely consistent with some central psychoanalytic concepts. Fenichel (1946) has expressed this in the following ways: "Character traits are the precipitates of instinctual conflicts. . . . In character attitudes, conflicts between impulses and fears may be relatively frozen. . . . This then is the situation. Once there was a conflict. The individual withdrew from this struggle by means of a permanent ego alteration. The forces that at one time opposed each other are now wasted in the useless and rigid defensive attitudes of the ego; the conflict has become latent." What this means, it would seem, is that a conflict between two or more emotions produces a new personality trait or character attitude which persists in time.

This raises the very interesting research question of what specific conditions of life bring about the mixing of emotions and the development of personality traits. Here too, the psychoanalytic literature has many suggestive observations. To cite just a few taken from Fenichel: "Stubborness is a passive type of aggressiveness, developed where activity is impossible. Compulsion neurotics are frequently particularly polite, accommodating and considerate. These are reaction formations that oppose aggressive tendencies. . . . Exceptional oral deprivation determines a pessimistic (depressive) or sadistic attitude." Fenichel suggests four general factors which determine the fixation of character attitudes: (*a*) The nature of the instinctual impulses that have to be warded off. (*b*) The time when the decisive conflict was experienced. (*c*) The content and intensity of the frustrations and the nature of the frustrating factors. (*d*) Whether substitute gratifications are available.

These concepts are, interestingly enough, closely related to some of those presented in connection with the frustration-aggression hypothesis (Berkowitz, 1958). It should also be possible, in time, to incorporate many of the concepts concerning conflict which have been developed through laboratory studies of animals and humans. Thus, there are many research possibilities open in the study of emotion-mixing.

There is a second important implication of the theory which is reflected in Table 1. As we go from primary to tertiary dyads, we go from what might be called the normal emotions of everyday life to the clinical emotions, that is, those attitudes toward people which frequently involve a person in psychiatric treatment. Thus, compare pride, love, curiosity, alarm, shame, misery, cynicism, and aggression with resentment, hate, guilt, despair, resignation, anxiety, envy, morbidness, and fatalism. Dyads composed of every fourth emotion listed on the emotion circle would be mixtures of opposites, and as already noted, opposites neutralize, immobilize, or inhibit. In general, the more opposed two emotions are the greater the conflict or immobilization produced by their mixture.

It should be evident that the theory can be extended to the mixtures of three emotions, although the problems of naming are proportionately greater. When we include triads, as well as dyads, and an intensity dimension as well, it becomes possible through the assumption of eight primary emotion dimensions to synthesize thousands of different combinations and thus potentially include all the emotions described in our language.

What are some other implications of the theory as outlined thus far? Take the question of a definition of emotion. The theory holds that an emotion is a total organismic reaction having physiological, behavioral, and sometimes subjective aspects. These are organized in some kind of pattern which differ from one emotion to another, or from one emotion-mixture to another. What the pattern is, in any given case, is an empirical question and can only be determined experimentally by a series of successive approximations. This implies too, that emotions can not be clearly separated from other total body reactions due to drugs, exercise, or fatigue. Many authors, including Freud, have noted the close interrelations between the effects of illness, fatigue, menstruation, and change of life, and emotions, implying that the patterns overlap, mix, and modify one another.

This view about the definition of emotion is not new or novel. Wenger (1950) presents essentially the same position in a paper, one in which he also quotes Dunlap writing in 1916 that "we must follow Lange, rather than James, in making the emotion not a *result* of the bodily state, but the bodily state itself." This view is related to that espoused by Babkin (1950), a physiologist who indicates that emotional behavior is based on an unconditioned reflex and is shown by widespread nervous reactions which involve many organs. Cannon (1939) describes rage as a "complex attitude which we do not have to learn. . . . It is a constant and uniform type of behavior, having features common in widely scattered races of men and even in lower animals. . . . (It has) the characteristics of a simple reflex, such as sneezing or coughing." Fenichel implies a similar view when he writes: "Affects are originally archiac discharge syndromes that supplant voluntary actions under certain exciting conditions." Such relative unanimity, of course, does not as such justify a definition; it simply indicates that the definition is not inconsistent with a considerable body of thought.

Another important question can be commented on in the light of this definition, and that is the problem of repression and control of emotions. From the point of view of the theory, one method for the control of emotions is through the mixing of emotions. We can control anger, for example, by mixing fear, or expectancy to either "dilute" the anger (the term is used metaphorically) or to partially inhibit it. Think, for example, of the fact that in ordinary conversation we often say to a person, "keep a stiff upper lip" or "keep your chin up" to hide or overcome sadness. Someone who has learned to hide all his emotions becomes a"dead pan" or "poker face." Do not these expressions indicate that we pit one bodily attitude against another in an effort to modify or reduce emotion?

Freud (1915) has pointed out that "the true aim of repression is to suppress the development of affect," and that "repression demands a constant expenditure of energy" (1915). How else is this to be interpreted except in terms of a pitting of one bodily attitude or set against another?

Several interesting speculations and bases for extension of the theory are implicit in the views described.

The first speculation concerns the function of psychotherapy; it is usually agreed that it is to help the individual recognize and resolve his conflicts. In the light of the considerations presented earlier concerning the mixing of emotions, is it not possible to restate this aim in the following terms: One purpose of psychotherapy is to help the individual recognize (become aware of, have insight into) the different emotions which are conflicting within him. This implies that "repressed anger," for example, must be a mixed emotion, or else, how would the anger be repressed. Doesn't the patient in therapy discover that his resentment is compounded of fear as well as anger and sadness? In a sense one function of psychotherapy is to encourage the individual to go through a process of mental analysis of his emotions in search of the basic components. The theory here presented, as it develops, should enable the therapist to be more keenly aware of this introspective analytic process.

This leads to a second speculation. One of the characteristics we find in all maladjusted individuals, is a restriction or narrowing of the range of emotional responsiveness. Such individuals may be consistently anxious, or depressed, or guilty, or resentful; or on the other hand, as in the case of many delinquents or psychopaths, aggressive, irritable, scornful, or cynical. At the same time, such individuals rarely, if ever, can openly experience and express any other kinds of emotions. The anxious or depressed person may be unable to express his joy or his anger; the delinquent may be unable to express his anxiety or his sadness at the death of another person. One sign of increasing mental health and adjustment is an increase in the flexibility of emotional responsiveness, and a decrease in the number of strongly conflicting emotions, that is, those that are most nearly opposite on the emotion circle. This would also imply a decrease in the preparatory muscular tensions which are associated with conflict, a point which is justified in an earlier paper (Plutchik, 1954). Thus, it seems that it might be possible to develop certain tests of what might be called "emotional responsiveness," as well as tests of muscular tension which could be used as indices of general emotional adjustment.

The theory is still a tentative and evolving one, and will undoubtedly undergo some revisions as new research data become available. At the same time, many problems of research are pointed to, for example, further studies of the behaviorial, physiological and subjective differences between the primary emotions, the development of an "emotional responsiveness" test as an index of adjustment, the development of standardized rules for naming emotion mixtures, the determination of the relation between personality typologies and the primary emotions, differences between psychosomatic patient groups in terms of emotion tendencies, etc. Particularly helpful may be the attempt to relate this emotion theory to a theory of personality.

References

1. Arnold, M. An excitatory theory of emotion. Chapter 2 in *Feelings and Emotions.* (Edited by M. L. Reymert.) New York: McGraw-Hill, 1950.
2. Ax, A. F. The physiological differentiation of fear and anger in humans. *Psychosom. Med.*, 1953, **15**, 433-442.
3. Babkin, B. P. The conditioning of emotions. Chapter 3 in *Feelings and Emotions.* (Edited by M. L. Reymert.) New York: McGraw-Hill, 1950.
4. Berkowitz, L. The expression and reduction of hostility. *Psychol. Bull.*, 1958, **55**, 257-283.
5. Bull, N., & Gidro-Frank, L. Emotions induced and studied in hypnotic subjects: II. *J. Nerv. & Ment. Dis.*, 1950, **112**, 97-120.
6. Bull, N. Towards a clarification of the concept of emotion. *Psychosom. Med.*, 1945, **1**, 210-214.
7. ———. The attitude theory of emotion. *Internat. Rec. Med.*, 1952, **165**, 216-220.
8. ———. The Attitude Theory of Emotion. New York: *Nerv. & Ment. Dis. Monog.*, 1951.
9. Cannon, W. B. The James-Lange theory of emotion. *Amer. J. Psychol.*, 1927, **39**, 106-124.
10. ———. Bodily Changes in Pain, Hunger, Fear and Rage. New York: Appleton-Century-Crofts, 1939.
11. Conant, J. B. On Understanding Science. New Haven: Yale Univ. Press, 1947.
12. Davis, R. C., Buchwald, A. M., & Frankman, R. W. Autonomic and muscular responses, and their relation to simple stimuli. *Psychol. Monog.*, 1955, **69**. (Whole No. 405).
13. Duffy, E. The psychological significance of the concept of "arousal" or "activation." *Psychol. Rev.*, 1957, **64**, 265-275.
14. Fenichel, O. The Psychoanalytic Theory of Neurosis. London: Rontledge & Kegan Paul, 1946.
15. Freud, S. The unconscious (1915). In *Collected Papers,* Vol. 4. London: Hogarth, 1925.
16. ———. Repression (1915). In *Collected Papers,* Vol. 4. London: Hogarth, 1925.
17. ———. Formulations regarding the two principles in mental functioning. In *Collected Papers,* Vol. 4. London: Hogarth, 1925.
18. ———. A General Introduction to Psychoanalysis. New York: Permabooks, 1953.
19. Funkenstein, D. H. The physiology of fear and anger. *Sci. Amer.*, 1955, **192**, 74-80.
20. Hebb, D. O. The Organization of Behavior. New York: Wiley, 1949.
21. Hill, D. C. The communication of emotions by facial expressions in the light of a new theory. M. A. Thesis, Hofstra College. Hempstead, N. Y., 1955.

22. James, W. What is emotion? *Mind,* 1884, **19**, 188-205.
23. Leeper, R. W. A motivational theory of emotion to replace "emotion as disorganized response." *Psychol. Rev.,* 1948, **55**, 5-21.
24. Lindsley, D. B. Emotion. Chapter 14 in *Handbook of Experimental Psychology.* (Edited by S. S. Stevens.) New York: Wiley, 1951.
25. Malmo, R. B. Anxiety and behavioral arousal. *Psychol. Rev.,* 1957, **64**, 276-287.
26. Masserman, J. H. A biodynamic psychoanalytic approach to the problems of feeling and emotion. Chapter 4 in *Feelings and Emotions.* (Edited by M. L. Reymert.) New York: McGraw-Hill, 1950.
27. Pasquarelli, B., & Bull, N. Experimental investigation of the mind-body continuum in affective states. *J. Nerv. & Ment. Dis.,* 1951, **113**, 512-521.
28. Plutchik, R. The role of muscular tension in maladustment. *J. Gen. Psychol.,* 1954, **50**, 45 ff.
29. ———. Further comments on the definition of emotion. *Psychol. Rec.,* 1957, **7**, 123-124.
30. ———. Some problems for a theory of emotion. *Psychosom. Med.,* 1955, **17**, 306-310.
31. Rapaport, D. Emotions and Memory. New York: International Univ. Press, 1950.
32. Schlosberg, H. Three dimensions of emotion. *Psychol. Rev.,* 1954, **61**, 81-88.
33. Wenger, M. A. Emotion as visceral action: An extension of Lange's theory. Chapter 1 in *Feelings and Emotions.* (Edited by M. L. Reymert.) New York: McGraw-Hill, 1950.
34. Wenger, M. A., Jones, F. N., & Jones, M. H. Physiological Psychology. New York: Holt, 1956.
35. Wolff, H. G. Life situations, emotions, and bodily disease. Chapter 24 in *Feelings and Emotions.* (Edited by M. L. Reymert.) New York: McGraw-Hill, 1950.
36. Woodworth, R. S. Experimental Psychology. New York: Holt, 1938.
37. Woodworth, R. S., & Schlosberg, H. Experimental Psychology. (2nd Ed.) New York: Holt, 1954.

Department of Psychology
Hofstra College
Hempstead, New York

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 173-177.

TIME ESTIMATES AS AFFECTED BY NEED TENSION AND RATE OF PROGRESS*

Department of Psychology, Trinity College, Hartford

ROBERT D. MEADE[1]

A. PROBLEM

Rosenzweig and Koht (5) reported that when *Ss* are under a condition of high need tension, time estimates are shorter than when under low need tension. High and low need tension conditions were essentially the same as ego and task involvement as described by Iverson and Reuder (2). However, Meade (4) has indicated that effects of need tension were confounded with those of perceived progress which he (3) and Hindle (1) have found to be inversely related to time estimates.

In re-tests of the Rosenzweig and Koht proposition with two different experimental designs, Meade (4) found no significant differences in time estimates between high and low need tensions. He did find, as he expected, that *Ss'* estimates were shorter in the second of two similar tasks occupying the same acount of time. However, in these studies, as well as the earlier one of Rosenzweig and Koht, no attempt was made to control rate of progress through the experimental tasks. The present study was done to test effects of both need tension and rate of progress on time estimates.

B. METHOD

1. *Subjects*

Forty-eight male college students, 12 in each of four experimental conditions were used.

2. *Apparatus*

Wooden block puzzles similar to those used by Rosenzweig and Koht and Meade, but contained 20 instead of four pieces, were used.

---

* Received in the Editorial Office on May 19, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a research grant, M-3193, from the National Institute of Mental Health, Public Health Service. The author wishes to thank Allen Schneider and Philip Newman who assisted in collecting the data.

3. *Procedure*

A factorial experimental design with two rates of progress and two levels of need tension was used and each *S* served in only one condition.

All *S*s were naive concerning the purpose of the experiment and had no objective means of estimating time. They were all given a standard 5 min. period with which they could later compare the time spent in the experimental task. During this time, they were asked to read a passage from a technical book with the belief that its content was to be remembered and was important for what followed. While *S* was reading, *E* left the room and returned at the end of 5 min. To this point, all *S*s were treated in the same manner.

The two levels of need tension and two rates of progress were brought about by instructions before the experimental task began and during the course of this task. *S*s in the low need tension groups were given the following instructions when *E* returned to the laboratory room:

> I have a series of tasks I want you to try out for practice. I want to see whether these tasks are suitable for a study I am planning to do in a few weeks. It may be that I will be able to use these tasks or I may have to discard them and you can help me find out. Your reactions will be greatly appreciated. These are for practice and you can take your time —when I've gotten the information I need, I will stop you and we will go on to something else.
>
> This is the first task with which I would like you to work. You are to manipulate these pieces of the puzzle in one and only one way to form a perfect square. This square will have no spaces or gaps in it but will be perfectly solid. In order to give you some indication of when you are on the right track, I will say "right" each time you fit two pieces together correctly; otherwise, I will say nothing.

*S*s in the high need tension groups were given these instructions:

> I am going to give you a new intelligence test. This test has already been proven to be one of the most accurate and reliable examinations yet developed. I am giving you the test for two reasons: first, we are interested in learning something about the general intelligence of college students and secondly, one of the faculty members here is doing a research project and we have to know how bright you are to determine if you are intellectually capable of participating in this project. In addition, it should be of interest to you since it will give you an excellent indication of how intelligent you are.
>
> This test is made up of several sections and naturally, you will have to do your very best on each section for me to assess your intelligence accurately. The first section is concerned with your problem-solving and

> spatial relations capabilities. You are to fit these wooden pices together in one and only one way to form a perfect square. This square will have no spaces or gaps in it, but will be perfectly solid. In order to give you some indication when you are on the right track, when you fit two pieces together correctly, I will say "right", otherwise, I will say nothing.

*S*s in the progress groups were told they had made a correct move 10 different times during the experimental session which lasted 5 min. *S*s in the non-progress groups were told nothing during this time indicating they had made no progress. *S* worked on the puzzle for 5 min. and at the end of this period, he was asked to compare the time spent on the puzzle with that spent reading and to determine which seemed longer to him. He was then told that the reading period was 5 min. long and was asked to estimate how much longer or shorter the puzzle session was. This estimate, subtracted or added to the 5 min. spent reading, was the dependent variable.

C. RESULTS AND DISCUSSION

Mean time estimates for the four experimental conditions are found in Table 1. Analysis of variance indicated that the rate of progress variable resulted in significant variation at less than the .05 level of confidence. ($F = 6.11$) while neither the need tension variable nor its interaction with rate of progress was significant ($F = .29$ and 1.29 respectively). Data from the post-experimental interview showed that 22 *S*s in the progress groups indicated they had experienced progress while 21 in the non-progress groups indicated they had experienced no progress.

TABLE 1

TIME ESTIMATES (IN MINUTES) AS AFFECTED BY NEED TENSION AND RATE OF PROGRESS

	Need tension	
	High	Low
Progress	3.4	4.6
No progress	5.5	5.1

The possibility that interrupting *S*s in the progress groups rather than progress information itself was responsible for the lower time estimates was checked with control groups and found to be insignificant.

It is proposed that the reason for failure to find significant differences between high and low need tension conditions was that the difference between these two conditions represented no quantitative difference in motivation to reach the goal. There was, however, the distinct possibility that there were qualitative differences in motivation between the two conditions of need tension. It can be argued, for example, that in the low need tension condition, *S* was motivated to finish the task because of its inherent interest value, to please *E*, to accept the challenge to his ingenuity, etc. Meade (3) has found that rate of progress affects time estimates when there is motivation to reach a goal.

If the proposition that there was motivation to reach goals in both the high and low need tension conditions is accepted, it is believed that the relationship between motivation, perceived progress, and time estimation holds for motivational conditions other than escape from boring and pointless work (used earlier by Meade). Such a generality would fit the theory that time estimates are conditioned by information which *S* receives concerning progress he is making toward ends he is motivated to attain.

D. Summary

Four groups of male college students were employed in a factorial experiment in which time estimates were made of the apparent length of a 5 min. period. Half of the *S*s were told their intelligence was being measured (high need tension) by the puzzle on which they worked while the other half was told to work on the puzzle for practice (low need). Both groups were told that they would be told each time they had made a correct move in the puzzle. By pre-arrangement, half of the *S*s in both groups was allowed to make progress while the other half was not.

There was an inverse relationship between progress and time estimates but no relationship between need tension and time estimates. The interaction between the two independent variables was also insignificant.

It was proposed that *S*s in *both* the high and low need conditions were motivated but toward different ends. Previous studies have established that when there is motivation to escape boring and pointless work, time estimates are inversely related to rate of progress. The present results are taken as increasing the generality of the kinds of motivation to which this applies and are interpreted in the general framework of theoretical considerations by Meade.

References

1. Hindle, H. Time estimates as a function of distance traveled and relative clarity of a goal. *J. Personal.*, 1951, **19**, 483-501.
2. Iverson, M., & Reuder, M. Ego-involvement as an experimental variable. *Psychol. Rep.*, 1956, **2**, 147-181.
3. Meade, R. D. Time estimates as affected by motivational level, goal distance and rate of progress. *J. Exp. Psychol.*, 1959, **58**, 275-279.

4. ———. Time perceptions as affected by need tension. *J. of Psychol.*, 1960, **49**, 249-253.
5. Rosenzweig, S., & Koht, A. The experience of duration as affected by need tension. *J. Exp. Psychol.*, 1933, **16**, 745-775.

Department of Psychology
Trinity College
Hartford 6, Connecticut

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 179-192.

THE PERCEPTION OF SHAPE FROM BINOCULAR DISPARITY CUES*

U.S. Army Medical Research Laboratory

WALTER C. GOGEL[1]

A. PROBLEM

This study will investigate some aspects of the perceived shape of binocularly observed three-dimensional objects as a function of their distances from *O*. When a three-dimensional object of constant physical size is placed at different distances from *O*, its image on the two retinas changes in a number of ways. One such change is a change in the binocular disparity associated with its depth component. This is illustrated in Figure 1 which can be interpreted as representing a top view drawing of two binocularly viewed three-dimensional objects. One of these objects has a frontal extent S_e and a depth extent X_{ef}. The other has a frontal extent S_g and a depth extent X_{gh}. The eyes of *O* are represented on the left with the distance between the pupils of the eyes labeled "A." In the lower portion of Figure 1, the angular subtense of the interpupillary distance *A* at the distances D_e, D_f, D_g, and D_h is shown as α_e, α_f, α_g, and α_h. The binocular disparity associated with X_{ef} is $\alpha_e - \alpha_f$ and the binocular disparity associated with X_{gh} is $\alpha_g - \alpha_h$. The physical shape of each of these three-dimensional objects is defined as the ratio of the physical depth *X* to the physical frontal extent *S*. Similarly, the perceived shape will be defined as the ratio of the perceived depth associated with *X* to the perceived frontal extent associated with *S*. If the same three-dimensional object is used at D_g as at D_e, then $S_e = S_g$ and $X_{ef} = X_{gh}$. But when $X_{ef} = X_{gh}$, $\alpha_e - \alpha_f \neq \alpha_g - \alpha_h$. This can be readily demonstrated by considering the relation of binocular disparity to linear physical distance. This relation is

$$\alpha_e - \alpha_f = \frac{A\ X_{ef}}{D_e^2 + D_e\ X_{ef}} \qquad (1),$$

with a similar equation applying to $\alpha_g - \alpha_h$, etc. Therefore, a three-dimensional object of constant physical size will produce different binocular

---

[*] Received in the Editorial Office on May 23, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank Richard E. Granda for his help in collecting and analyzing the data of this study.

disparities when placed at different distances from *O*. The problem is to determine what consequences this has for the perception of three-dimensional shape.

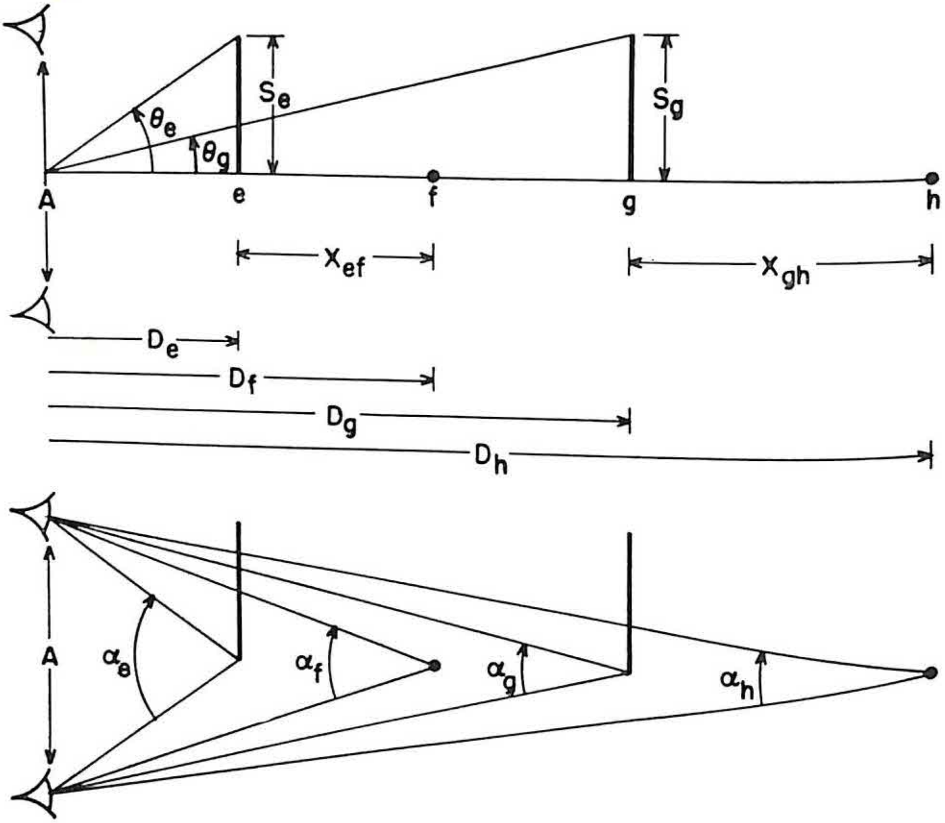

FIGURE 1
A Schematic Diagram for Considering Three-Dimensional Binocular Judgments

It is clear that if the effect of a changing binocular disparity upon the perception of shape is to be tested, only binocular cues should determine the perceived extent associated with the depth component. This means that only binocular cues should perceptually relate *f* to *e* or *h* to *g* in Figure 1. This was the only restriction with respect to the available cues of distance which was placed on the study. As will be discussed later, any number of cues to the perceived absolute distance of the three-dimensional object (the perception of its distance from *O*) can be present. It will be demonstrated that, under these conditions, the perceived shape of a three-dimensional object changes as a function of distance in an orderly manner. It will be demonstrated also that these changes in perceived shape are accurately predicted from a theoretical position which stresses the relational characteristics of retinal stimuli as determiners of perception.

B. METHOD

The general method used in this experiment will be discussed with the aid of Figure 1. As before, consider that the drawings of Figure 1 represent three-dimensional objects with the three-dimensional object at D_e being defined by S_e and X_{ef} and the three-dimensional object at D_g being defined by S_g and X_{gh}. Let $S_e = S_g$, but instead of X_{ef} necessarily being physically equal to X_{gh} (as in the example given above), suppose that each of these depth intervals are adjustable by *O*. In the experiment, the task of *O* was to adjust *f* in depth until the distance X_{ef} was perceptually equal to S_e and to adjust *h* in depth until X_{gh} was perceptually equal to S_g, etc. Thus, physically equal frontal extents were placed at several distances and *O* adjusted the depth interval until, in each instance, the depth extent was perceptually equal to the frontal extent. It is clear that *O* was producing perceptually equal three-dimensional shapes at the several distances. The experimental problem is to determine how these adjusted depth intervals (*X*) varied with changes in distance (*D*).

1. *Apparatus*

a. The visual alley. A schematic perspective diagram of the visual alley used in the experiment is shown in Figure 2. The alley was enclosed by a black wall on the left and by black curtains on the right. It was illuminated by a series of overhead fluorescent lights. The floor of the alley was covered with tan cloth. A small rectangle of black cardboard (4 in. wide by 1 in. high) was presented on a stand at either 5.25, 10.50, or 15.75 feet from *O*. The rectangle was presented perpendicular to the median plane of *O*, with its lower edge 6¼ inches above the floor of the visual alley. A small luminous disc was adjustable in depth by O. It moved along a line-of-sight which passed ½ inch above the right edge of the rectangle for each of the distance positions of the rectangle. The three alternate positions of the rectangle and the path of depth movement of the disc are schematically illustrated in Figure 2. Relating Figure 2 to Figure 1, a top view of the rectangle would be represented by *e* or *g* and a top view of the disc by *f* or *h* in Figure 1.

b. Generating the disc. The movable disc was stereoscopically generated by using a pair of reflex sights, with one reflex sight before each eye of the subject. Each reflex sight contained a circular aperture source of light. The light from this source passed through an optical system and emerged as parallel light. This parallel light was then reflected from a reflecting-transmitting surface into an eye of *O* as a luminous disc. The reflecting-transmitting surface also permitted *O* to see objects which were actually located in front of himself. By turning an adjustment knob, the pair of reflex sights could be rotated simultaneously in opposite directions thereby changing in continuous but opposite amounts the lateral positions of the disc image in each eye. When *O* looked through the stereoscopic instrument at the black rectangle in the visual alley and binocularly fused the two disc images, the

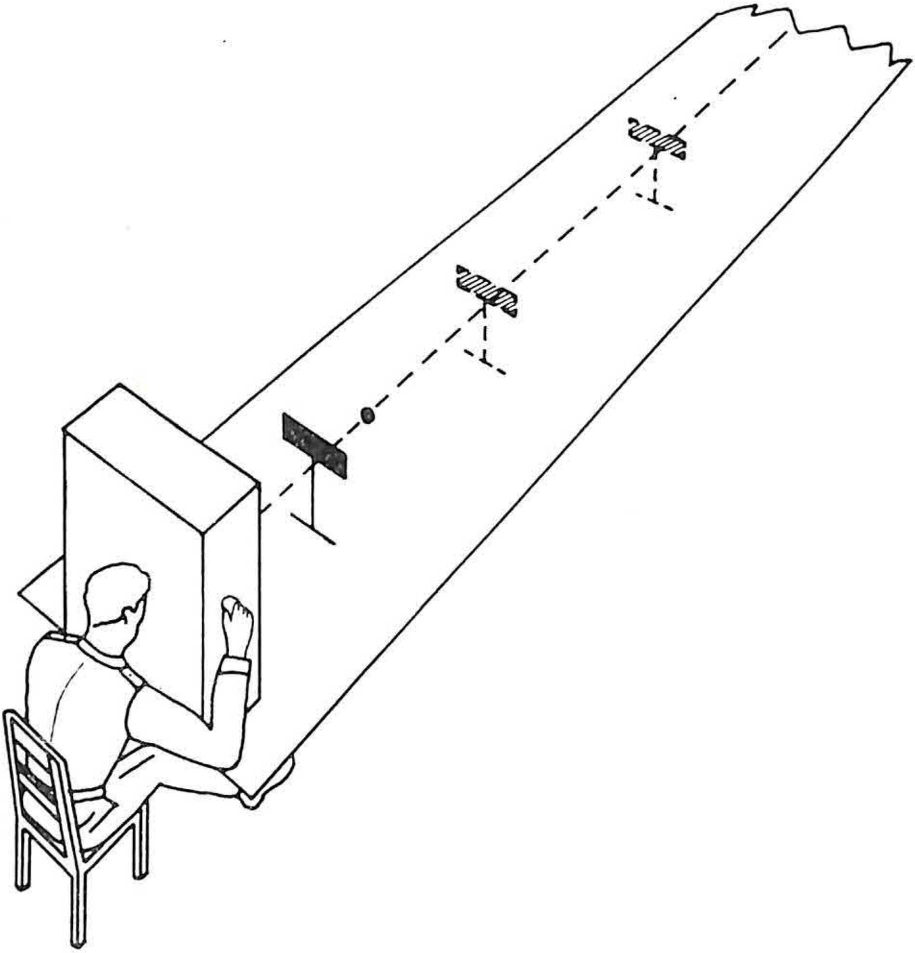

FIGURE 2
A Perspective Schematic Diagram of the Visual Alley

disc appeared to occupy a position in distance with respect to the rectangle. When *O* turned the adjustment knob, the lateral retinal positions of the two disc images changed and the disc appeared to move directly toward or away from *O* depending upon the direction of the knob rotation. The movable disc will be called the disc reticle. The binocular disparity measured in angles between different distance positions of the disc reticle was indicated by a dial attached to the stereoscopic instrument. The instrument used to generate the disc reticle is the same as that described by Harker (6).

c. Depth cues in the visual alley. The floor and walls of the alley were visible to *O*. The distance position of the black rectangle with respect to the floor of the alley was evident through its supporting stand. Thus, a variety of cues were present to determine the perceived absolute distance of the black rectangle. This was not the case with respect to the disc reticle. The disc appeared to be suspended above the floor of the alley without any physical support. Also, the path of depth movement of the disc was along a line of sight directly above the line of sight to the rectangle. This isolation of the disc from the floor of the alley but not from the rectangle enhanced the effectiveness of the binocular disparity cue between the disc and rectangle in determining the perceived position of the disc (4). The binocular disparity between the disc and rectangle was the determining cue for the perception of relative depth between these two objects.

2. *Procedure*

Eighteen *O*s made two types of adjustments of the depth position of the disc. In one adjustment called the equidistance adjustment *O* moved the disc until it appeared to be at the same distance as the right edge of the black rectangle. In the second type of adjustment called the depth-duplication adjustment *O* moved the disc as far behind the right edge of the rectangle as the left edge of the rectangle appeared to be to the left of the right edge. In other words, in the depth duplication adjustment, the subject duplicated the apparent width of the rectangle in apparent depth. The difference between an equidistance adjustment and an immediately following depth duplication adjustment constituted a single score. By using the calibration constant of the instrument and the interpupillary distance of *O*, this score could be converted to either a binocular disparity ($\alpha_e - \alpha_f$) or a linear depth extent (X_{ef}). Three such scores were obtained for each *O* at each of the three distance positions of the rectangle. Following this, *O* estimated the distance of the rectangle from himself in feet and its width in inches.

C. RESULTS

The average depth interval which appeared equal to the width of the rectangle at the three distances is given in Table 1 both as binocular disparity in seconds of arc and linear distance in centimeters. The means and standard deviations of Table 1 were found from the distributions of 18 scores, one from each subject, with each score being the average difference between the three equidistance and three depth duplication adjustments for each distance position of the rectangle. By considering the rectangle and the adjusted depth component as constituting a three-dimensional object, the centimeter results shown in Table 1 can be simply stated. The number of centimeters of physical depth required in order for the depth component of the three-dimensional object to appear equal to the frontal component increased linearly with distance. This is indicated in Figure 3 in which the centimeter data in Table 1 are seen to satisfy the relation that the required *X* is a direct linear function of *D*. Conversely, the binocular disparity data of Table 1 satisfy the relation that the required $\alpha_c - \alpha_f$ is inversely related to *D*.

TABLE 1
RESULTS FROM DUPLICATING A 10.2 CM. FRONTAL EXTENT IN APPARENT DEPTH, AT THREE DISTANCES, WITH BINOCULAR DISPARITY CUES (N = 18)

Distance of rectangle (D)	Binocular disparity in sec. Mean	*SD*	Linear depth (X) in cm. Mean	*SD*
Near (5.25 ft)	780	360	17	9
Medium (10.50 ft)	390	170	33	16
Far (15.75 ft)	240	100	45	20

Correlations between the binocular disparity results from the several distances were calculated. The *t* test was used to determine the statistical significance of the differences between the average binocular disparities of Table 1. The resulting *r* and *t* values involving the near (N), middle (M) and far (F) distance conditions were: $r_{NM} = .75$, $t_{N\text{-}M} = 6.77$; $r_{NF} = .86$, $t_{N\text{-}F} = 12.54$; $r_{MF} = .85$, $t_{M\text{-}F} = 3.69$. It is evident from the large standard deviations in Table 1 that the *O*s differed in their depth adjustments. But, from the values of *r*, it is evident also that there was individual consistency in the depth adjustments which remained despite changes in *D*. This is in agreement with the existence of an individual observer constant and such a constant (C) will be described later in this paper. Each of the above *t* values was beyond the .01 level of confidence. Therefore, it is clear that the average depth adjustments which were completed at the different distances were not the same, i.e., the changes in the mean results of Table 1 as a function of distance (D) are real.

The average verbal reports of the distance to the rectangle was 3.6, 7.3, and 10.9 feet and the average verbal report of the width of the rectangle was 3.9, 4.0, and 3.9 inches for the 5.25, 10.5, and 15.75 feet positions respectively. The verbal reports of distance show the usual tendency to underestimate the absolute physical distance in situations of this sort (5). However, all subjects responded with an increase in estimated distance whenever the physical distance of the rectangle from *O* was increased. It seems that effective cues to the absolute distance of the rectangle were present. Also, it

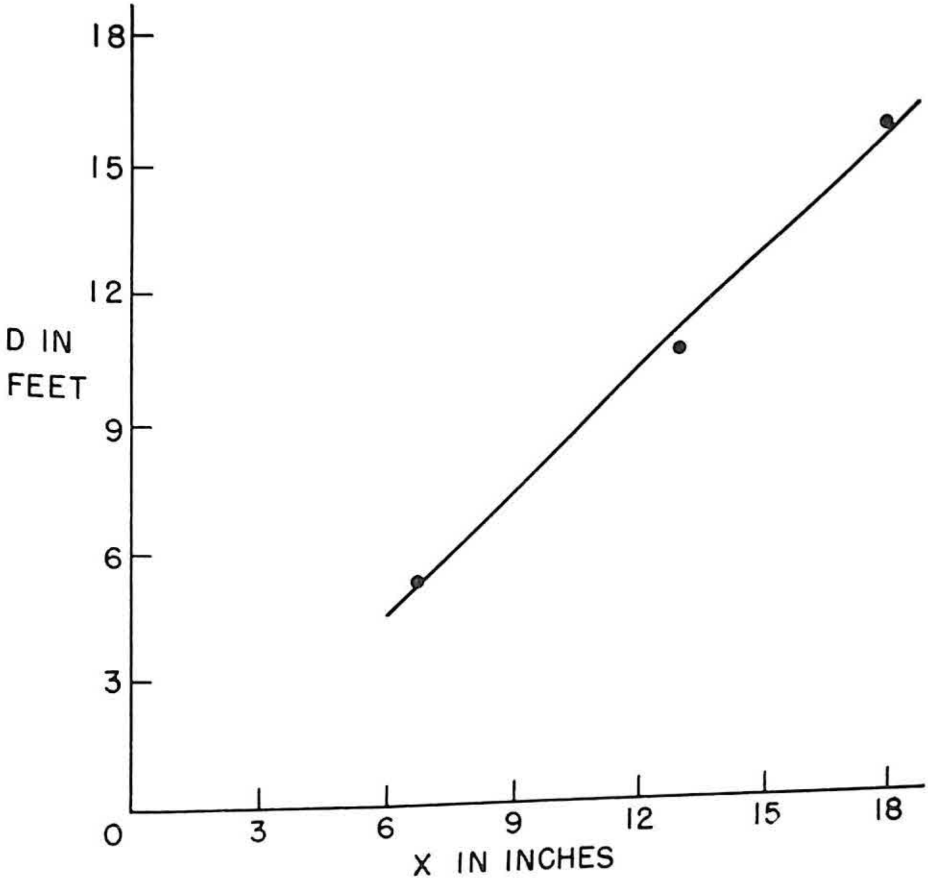

FIGURE 3
THE AVERAGE DEPTH INTERVALS (X) AT DIFFERENT DISTANCES (D) REQUIRED TO DUPLICATE 10.2 CM. OF FRONTAL EXTENT IN APPARENT DEPTH

will be noted that very similar average estimates of the width of the rectangle were obtained for all distances of the rectangle. Evidently, practically perfect frontal size constancy occurred in this experiment even though, from the results of Table 1, perfect three-dimensional shape constancy was not present.

D. DISCUSSION

1. *General Theory*

The results obtained in this study can be understood by using a point-of-view which has been introduced in a previous article (3). This point-of-view emphasizes the perceptual interrelation of frontal and depth extents and postulates that the perception of absolute distance is not necessary in order for this interrelation to occur. The application of this to the present study can be illustrated by Figure 1. In Figure 1 consider the three-dimensional object specified by the frontal extent S_e and the depth extent X_{ef}. On the retina, the angular size resulting from S_e is θ_e and the binocular disparity resulting from X_{ef} is $a_e - a_f$. For didactic purposes consider the case in which the three-dimensional object is quite small. In this case both S_e and X_{ef} will be small and $a_e - a_f$ can be written as Δa. Let the perceived or estimated size associated with S_e be labeled E_e and the perceived depth associated with Δa be labeled $\Delta a'$. The hypothesis which will be applied to the present study is that

$$\frac{\Delta a'}{E_e} = \frac{1}{C}\,\frac{\Delta a}{\theta_e} \tag{2}$$

where C is an individual constant. Equation 2 states that the ratio of the perceived depth $\Delta a'$ to the perceived frontal width E of the three-dimensional object is proportional to the ratio of the angular representation of these two extents on the eye. In other words, for small three-dimensional objects, it is hypothesized that the perceived ratios of the several extents is proportional to their retinal ratios with $\frac{1}{C}$ being the constant of proportionality. It will be recalled in this connection that Δa is the retinal representation of the depth extent X_{ef}, and θ_e is the retinal representation of the frontal extent S_e.

Equation 2 represents the hypothesis of this study. It will be noted that it contains no terms for perceived absolute distance. This means that Equation 2 is expected to apply regardless of whether the small three-dimensional object appears to be near, far, or at an indefinite distance. It is this quality which allows Equation 2 (and the equations which will be derived from it) to be appropriate for situations like those of the present experiment where many cues to the absolute distance of the frontal rectangle were present. However, it should also be noted that only binocular disparity cues are allowed within the interval specified by Δa, i.e., the interval X_{ef}. If this is not the case, equations which will be derived from Equation 2 are not expected to apply. Indeed this has been demonstrated by an experiment in which cues in addition to binocular disparity were deliberately introduced within this interval. Under these latter circumstances, as expected, equations derived from Equation 2 were found to apply with respect to the direction of the results but not with respect to magnitude (1).

It was pointed out that Equation 2 applied only to small three-dimensional objects. But the perceptual addition of the apparent depth extents of a number of small objects can perceptually result in a larger object. In essence, this is the method which is used to derive useful equations from Equation 2. It will be shown that these equations quite accurately predict the change in data indicated in Figure 3.

If Equation 2 is written in terms of differentials rather than increments and is rearranged, it follows that

$$da' = \frac{E}{\theta}\,\frac{da}{C} \tag{3}$$

In Equation 3, the term $\frac{E}{\theta}$ is the perceived size of the frontal object per unit of its angular size when the object is located at a particular distance D. However, since this term is a ratio it is independent (within limits) of the physical extent of any particular frontal object. In other words, a particular value of $\frac{E}{\theta}$ rather than applying only to a particular size of frontal object applies to all portions of the frontal plane at the particular distance. The ratio $\frac{E_e}{\theta_e}$ is the perceived size per unit of angular frontal size of the frontal plane at the distance D_e. The ratio $\frac{E_v}{\theta_v}$ designates the perceived size per unit of angular frontal size of the frontal plane at any distance D_v. From Equation 3 it follows that

$$(a_e - a_f)' = \int_f^e da' = \frac{1}{C}\int_f^e \frac{E_v}{\theta_v}\,da \tag{4}$$

where $(a_e - a_f)'$ is the perceived depth associated with the binocular disparity

$\alpha_e - \alpha_f$. It will be noted in Equation 4 that $\frac{E_v}{\theta_v}$ can vary as a function of α. The term $\frac{1}{C}$ however is an individual constant and so is independent of α.

Equation 4, unlike Equations 2 and 3, applies to extents of any magnitude. But Equation 4 is not in a form which can be applied to the present experiment since it does not specify the perceived shape of a three-dimensional object. To remedy this deficiency let m_e be the perceived shape of the three-dimensional object at D_e, where m_e is defined as the ratio of the perceived depth of the three-dimensional object at D_e to its perceived frontal extent E_e, i.e.,

$$m_e = \frac{(\alpha_e - \alpha_f)'}{E_e} \tag{5}$$

Combining Equations 4 and 5 gives

$$m_e = \frac{(\alpha_e - \alpha_f)'}{E_e} = \frac{1}{CE_e} \int_f^e \frac{E_v}{\theta_v}\, d\alpha \tag{6}$$

which is an equation for the perceived shape of the three-dimensional object at D_e. Similarly, for the three-dimensional object at D_g (see Figure 1),

$$m_g = \frac{(\alpha_g - \alpha_h)'}{E_g} = \frac{1}{CE_g} \int_h^g \frac{E_v}{\theta_v}\, d\alpha \tag{7}$$

In order to apply Equations 6 and 7 to the present study these equations must be made more explicit by specifying the relation between $\frac{E_v}{\theta_v}$ and α. The relation between θ and α is easily found from the following equations in which θ and α are expressed in radians.

$$\alpha = \frac{A}{D} \tag{8}$$

and

$$\theta = \frac{S}{D} \tag{9}$$

Expressing the appropriate relation between E_v and α is more difficult. For this purpose consider a frontally presented object of constant frontal extent S which is located at different distances D_v. Three cases will be considered. Suppose that the perceived size E of the constant S varies inversely with distance (or, from Equation 8, directly with α). This is a case of zero amount of frontal size constancy and can be expressed as

$$\frac{E_v}{E_e} = \frac{D_e}{D_v} \tag{10}$$

where E_v is the perceived size of a constant S at any distance D_v and E_e is the perceived size of the same S at a particular distance D_e. Or, suppose that E remains constant despite changes in D (or α). This is a case of perfect frontal size constancy and can be expressed as

$$E_v = E_e \tag{11}$$

Equation 11 expresses the case in which the perceived frontal size (E_v) of an object at any distance is equal to its perceived frontal size (E_e) at a particular distance, i.e., E_v in this instance is a constant. Suppose, finally, that E increases directly with D (inversely with α). This would be a particular case of over frontal size constancy and can be expressed as

$$\frac{E_v}{E_e} = \frac{D_v}{D_e} \tag{12}$$

From Equations 10, 11, and 12 together with Equation 8 it is apparent that the relation between E and α (or D) will vary as a function of the amount of frontal size constancy present. Equations 10, 11, and 12 can all be expressed by the relation

$$\frac{E_v}{E_e} = \frac{D_e^n}{D_v^n} \tag{13}$$

In Equation 13, when $n = +1$, Equation 10 results; when $n = 0$, Equation 11 results, and when $n = -1$, Equation 12 results. Thus n in Equation 13 is a term for frontal size constancy. Zero amount of frontal size constancy is expressed by setting $n = +1$. Perfect frontal size constancy is expressed by setting $n = 0$ and an amount of over frontal size constancy is expressed by setting $n = -1$. It will be noted that Equation 13 is not theoretically derived. It is merely a way of noting how perceived frontal size will change with distance if various amounts of frontal size constancy are present. The value of n must be known before Equation 13 is useful. For example, in the present study the average reported width of the rectangle was practically constant at all distances. This means that n was essentially zero.

Equations 8, 9, and 13 can be used to transform Equations 6 and 7 to a useable form. From Equations 8, 9, and 13, Equations 6 and 7 become

$$m_e = \frac{A}{CSa_e{}^n} \int_f^e a_v{}^{n-1} da \qquad (14)$$

and

$$m_g = \frac{A}{CSa_g{}^n} \int_h^g a_v{}^{n-1} da \qquad (15)$$

But, in this study $m_e = m_g = 1$ and therefore,

$$\frac{1}{a_e{}^n} \int_f^e a_v{}^{n-1} da = \frac{1}{a_g{}^n} \int_h^g a_v{}^{n-1} da \qquad (16)$$

Simplifying Equation 16, it follows that for all values of n

$$\frac{a_e}{a_f} = \frac{a_g}{a_h} \qquad (17)$$

From Equation 8 and recalling that $X_{ef} = D_f - D_g$ and $X_{gh} = D_h - D_g$, Equation 17 becomes

$$\frac{X_{ef}}{X_{gh}} = \frac{D_e}{D_g} \qquad (18)$$

Or, using Equation 1, Equation 18 becomes

$$\frac{a_e - a_f}{a_g - a_h} = \frac{D_g}{D_e} \qquad (19)$$

The conclusion from Equation 18 is that the physical depth required in order to perceptually duplicate a frontal extent of constant size varies directly with the distance. It will be noted that this is exactly the result which is shown in Figure 3. The results from this experiment therefore indicate that Equations 2, 4, 6, and 7 are correct. Taken together, these equations specify the perceptual consequences of binocular disparity cues for a number of kinds of judgments (2). Additional consequences of these equations will be evaluated in future reports.

2. *The Observer Constant C*

In deriving Equation 16 from Equations 14 and 15 it was assumed that C was a constant for an individual O. The agreement of Equation 18 with the obtained data substantiates this assumption. But, the average value of C can be determined directly from the experimental results. From Equation 14, since $m = 1$ in this experiment, it follows that for this study,

$$C = \frac{A}{S{a_v}^n} \int_f^e {a_v}^{n-1} da \qquad (20)$$

But, the verbal reports indicate that in this experiment $n = O$. Substituting $n = O$ in Equation 20 and simplifying by means of Equations 8 and 9 results in

$$C = \frac{a_e}{\theta_e} \ln \frac{a_e}{a_f} \qquad (21)$$

Equation 21 was applied to the average results from the present study for each of the three distances used. It was found that *C* was .063, .062, and .057 for the 5.25, 10.50, and 15.75 foot distances respectively. The similarity of these values to each other supports the hypothesis that *C* is independent of a (or *D*). Referring to Equation 2, an interpretation of these computed values of *C* is that for this group of *O*s, on the average, a binocular disparity whose angular value was about 6 per cent of the frontal angular width would be required in order for a small three-dimensional object to be perceived as a cube regardless of its distance from *O*.

E. Summary and Conclusions

In this study, shape judgments were investigated with respect to a three-dimensional object defined as a frontal component and an adjustable depth component. Eighteen *O*s adjusted the depth component to appear equal to the frontal (width) component at three distances. A requirement of the study was that only binocular disparity cues support the depth component whereas the apparent distance position of this three-dimensional object was supported by a number of cues. It was found that the magnitude of the depth component which was required in order to appear equal to the frontal width increased linearly with distance. This result was predicted from a theoretical position which emphasizes the significance of the relational characteristics of retinal stimuli. In agreement with this point-of-view, an observer constant was postulated and was found to be substantiated by the experimental results.

References

1. Gogel, W. C. The perception of shape from binocular disparity cues. *U.S. Army Med. Res. Lab.*, 1958, Report No. 331, 1-27.
2. ———. The perception of space with binocular disparity cues. *U.S. Army Med. Res. Lab.*, 1959, Report No. 379, 1-24.

3. ———. Perceived frontal size as a determiner of perceived binocular depth. *J. of Psychol.*, 1960, **50**, 119-131.

4. GOGEL, W. C., & HARKER, G. S. The effectiveness of size cues to relative distance as a function of lateral visual separation. *J. Exper. Psychol.*, 1955, **50**, 309-315.

5. GOGEL, W. C., HARTMAN, B. O., & HARKER, G. S. The retinal size of a familiar object as a determiner of apparent distance. *Psychol. Monog.*, 1957, **71**, No. 13 (Whole No. 442).

6. HARKER, G. S. Instrument for the study of binocular vision. *J. Opt. Soc. Amer.*, 1955, **45**, 826-828.

Psychology Division
U.S. Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 193-195.

RELATIONSHIP BETWEEN *PGR* AND MEASURES OF EXTRAVERSION, ASCENDANCE, AND NEUROTICISM*[1]

Payne Whitney Psychiatric Clinic, New York Hospital

ARLINE BRONZAFT, ROSLYN HAYES, LIVINGSTON WELCH,
AND MYRON KOLTUV

A. INTRODUCTION

PGR has frequently been used to validate other personality measures and clinical observations. Jones (1) correlated *PGR* with personality traits in a stress situation and found a significant relationship between extraversion and *PGR*. Previous studies in our laboratories (2, 3), investigating differences in *PGR* between patients in various psychiatric categories and controls, have demonstrated differences in a simple avoidance situation. Rockwell (3) studied *PGR*s of psychoneurotic patients and hypothesized that the defensive reaction of a psychoneurotic is manifested in lowered *PGR* reactions.

In the present study the following hypothesis was tested: *PGR* in a simple avoidance situation is related to measures of extraversion, ascendance, and neuroticism.

B. PROCEDURE

The subjects were 46 female college students between the ages of 17 and 32 years, enrolled in psychology courses at a large municipal college. Since participation was a course requirement, they were not volunteers.

The Maudsley Personality Inventory developed by Eysenck was used to measure introversion-extraversion and neuroticism. The Allport A-S Reaction Study was used to measure ascendance-submission.

PGR apparatus consisted of a custom-made pathometer and a Texas Instrument Rectiriter which gave a paper tape recording.

The two tests were administered to the subjects at different times from a few days to several months apart.

Subjects were tested individually on another occasion in the avoidance situation which was disguised as a mock reaction time experiment. The situation is as described in previous studies with the minor change in instructions that subjects were told to watch for the green light before each trial.

---

* Received in the Editorial Office on June 6, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] From the Department of Psychiatry, Cornell University Medical College (Payne Whitney Psychiatric Clinic), and the Department of Psychology, Hunter College, New York City.

Essentially the test consisted of asking the subject to press a lever at the green light signal for three trials. At the end of these trials, his reaction time was criticized for slowness and he was told to press faster on the fourth trial to avoid the sound of a Klaxon horn. Although reaction time was never measured, the horn was sounded for every fifth subject for the sake of credibility.

C. RESULTS

Means and SD's were computed for the subjects on the A-S scale and on the extraversion and neuroticism scales of the M. P. I. Individual *PGR*s to the critical instructions were measured as the ratio of the subject's peak response (in ohms) to his basal skin resistance prior to the critical instructions. The mean and SD of the *PGR*s were also computed (Table 1).

Table 2 presents the correlations between *PGR* and the measures of extraversion, ascendance, and neuroticism. None of these correlations were significant.

TABLE 1
MEANS AND SDs ON FOUR VARIABLES (N=46)

Variables	Mean	SD
Extraversion (MPI)	27.63	8.06
Neuroticism (MPI)	20.02	10.28
Ascendance (Allport)	1.22	17.98
PGR	.19	.23

TABLE 2
RELATIONSHIP BETWEEN PGR AND MEASURES OF EXTRAVERSION, ASCENDANCE, AND NEUROTICISM (N = 46)

Variables	Correlations
Extraversion—PGR	—.103
Neuroticism—PGR	.232
Ascendance—PGR	—.148

D. DISCUSSION

Jones measured *PGR* in a mild stress situation and differentiated between a high reactive group and a low reactive group. He then correlated *PGR* with various social traits. He found that the low reactive group was aggressive, dominating, irritable, easily excited, uninhibited, and less restrained in social behavior. The high reactive group was calm, more deliberate, more cooperative, more responsible, and more restrained in social behavior. Jones postulated that the low reactive group's more extraverted expressive pattern might be associated with restricted physiological reactions such as *PGR*. Jones says it is also possible that internalization of emotional tension may be more indicative of those who display socially approved behavior. In our group, the negative correlation of —.103 between *PGR* and extraversion gives some little support to Jones' hypothesis, but the correlation is negligible and not significant.

This difference in results between Jones' study and this study may be due to the fact that different measures of extraversion-introversion were used. Studies in the area of personality often yield conflicting results because of the questionable validity or reliability of the testing instruments as well as differences in the test author's conception of the underlying trait.

Although Rockwell found that psychoneurotic patients gave the lower *PGR*s, our correlation of .232 between neuroticism and *PGR* was not significant. Our range of scores on neuroticism was necessarily more limited than Rockwell's because our subjects were college students and we had no subjects clinically diagnosed as neurotic. In addition, different *PGR* test situations and different measurements of neuroticism were used.

E. Summary and Conclusions

The Maudsley Personality Inventory and the Allport A-S Reaction Study were administered to 46 female college students. *PGR*s were also obtained in a simple avoidance situation. A relationship between *PGR* and measures of ascendance, neuroticism, and extraversion was hypothesized. No significant correlations were found between *PGR* and any of the measures. Previous studies have reported significant relationships between *PGR* and measures of extraversion-introversion and neuroticism, particularly where the population studied included subjects in various psychiatric categories. The difference between these results and those reported by other investigators may be due to differences in the tests used to measure *PGR*, extraversion, ascedance, and neuroticism, to the limitations of the sample studied, or to the unreliability of the tests.

References

1. Jones, H. E. The study of patterns of emotional expression. *In* M. L. Reymert (*Ed.*), *Feelings and Emotions: The Mooseheart Symposium*. New York: McGraw-Hill, 1950.
2. Koltuv, M., Hayes, R., Fuchs, A., & Welch, L. A comparison of the GSR's of psychiatric patients and normals in an avoidance situation. *J. of Psychol.*, 1959, **48**, 115-119.
3. Rockwell, F. V., Welch, L., Kubis, J., & Fisichelli, V. Changes in palmar skin resistance during the Rorschach test. *Mschr. Psychiat. & Neurol.*, 1957, **113**, 129-152.

Payne Whitney Psychiatric Clinic
New York Hospital
525 East 68th Street
New York 21, New York

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 197-204.

THE DEVELOPMENT AND EMPLOYMENT OF VTAT's OR PICTURELESS TAT's*

Child Guidance and Speech Correction Clinic, Jacksonville, Florida

DELL LEBO

A. INTRODUCTION

Odd as it may seem at first glance, a pictureless TAT offers several advantages. In the first place, a pictureless or verbal Thematic Apperception Test (VTAT) would seem to be more ambiguous than graphic representation. In an effort to cause the pictorial stimuli to retreat from consciousness and the underlying personality to be revealed more clearly Murray (1951) suggested that subjects examine a picture for 20 seconds or so and then lay it aside prior to responding. While Bellak (1952) described a method in which patients were to be shown a picture, then instructed to look at the ceiling and project the picture just seen onto it in imagination. Subjects were asked to describe this picture and tell the traditional story about it. Verbal descriptions of the TAT card presented visually (ViTAT) for subjects to read, or presented vocally (VoTAT) by being read to subjects, would eliminate contact with the cards and require subjects to construct mentally their own pictures from the word stimuli. A VTAT would seem to be the most ambiguous of TAT's insofar as stimulus ambiguity is concerned.

In forcing the client to rely more upon himself and less upon the stimuli the evasiveness of card description would also be eliminated. Subjects might read ViTAT stimuli aloud or speak VoTAT words aloud for clarity. Such performance, however, obviously would be a repetition and not an unskilled avoidance of a response.

Other important uses of a VTAT would be in group testing where the stimuli could be administered in recorded (VoTAT) or printed (ViTAT) form. Massive amounts of normative data, worthy of machine tabulation, could be collected easily. Such material has begun to appear recently in small quantities due to the difficulties inherent in several methods of administering group TAT's, whereas, the VTAT seems designed for large scale uses. Group use of a VTAT among different socioeconomic and age levels might quickly and easily establish standardization data of importance.

---

* Received in the Editorial Office on June 8, 1960, and published immediately at Provincetown, Massachusetts, by means of financial assistance from the American Foundation for the Blind. Copyright by The Journal Press.

Intracultural comparison would be simplified as well. It seems easier to translate VTAT stimuli into appropriate words than to have TAT pictures redrawn in conformity with surroundings. For example, suitable words for TAT card 12BG, "A row boat is drawn up on the bank of a woodland stream," are probably more readily available than suitable pictures for most cultures. Local flora are more easily and correctly imagined than depicted.

Projective testing with the blind could be advanced through the use of a VoTAT. While the number of suitable evaluative procedures for blind persons is increasing, there is a tendency to validate such new techniques by means of clinical impression rather than by comparison with older methods. Also, several such new methods are unsuitable for sighted persons or have no history of such application. VoTAT's would serve admirably as a standard for comparison. The more closely the VTAT resembled the TAT the more closely could the responses of the blind be compared with TAT norms and the more closely could other projective techniques developed for use with the blind be contrasted for effectiveness with an established personality examination. This matter has been discussed further by Lebo and Bruce (1960).

B. RANKIN'S VTAT

Information about the availability of the TAT for use with the sighted first appeared in 1935 (Morgan and Murray). It has been used in verbal form with the blind since 1952.[1] This use of the VTAT is by Rankin who is, herself, blind.[2] It is based upon her modifications of the TAT descriptions in Murray's *Manual* (Murray, 1943).

An examination of Rankin's VTAT[3] and the descriptions given by Murray reveal many differences. In view of the use of Murray's descriptions in the preliminary investigations of another development, by the present writer and his co-workers, both Murray's TAT descriptions and Rankin's VTAT stimuli are presented for comparison. To make comparison easier Murray's description is given first, followed by Rankin's. To make identification easier Murray's descriptions are preceded by an *M* and Rankin's stimuli by an *R*.

C. MURRAY'S VTAT AND RANKIN'S VTAT

M1. A young boy is contemplating a violin which rests on a table in front of him.

---

[1] Rankin, S. K. Personal communication. April 9, 1958.

[2] Maddox, J. K. Personal communication. February 5, 1958.

[3] Rankin, S. K. Personal communication. May 1, 1958.

R1. A boy is sitting at a table on which there is a violin and some papers which could be music.

M2. Country scene: in the foreground is a young woman with books in her hand; in the background a man is working in the fields and an older woman is looking on.

R2. This is a farm scene. There is a muscular man plowing with a team. In the foreground, there are two women. One appears to be pregnant. The other appears to be younger and has some books on her arm. The women are not looking at each other.

M3BM. On the floor against a couch is the huddled form of a boy with his head bowed on his right arm. Beside him on the floor is a revolver.

R3BM. Here is a figure collapsed in front of a cushioned bench. The figure's head rests on the bench; an arm circles the head. The other arm hangs off the edge of the seat, and an object which appears to be a gun lies on the ground nearby.

M3GF. A young woman is standing with downcast head, her face covered with her right hand. Her left arm is stretched forward against a wooden door.

R3GF. A woman with drooping head is standing in a doorway; there is darkness behind her. Her face is partly covered by one hand. The other arm is outstretched along a wooden door frame.

M4. A woman is clutching the shoulders of a man whose face and body are averted as if he were trying to pull away from her.

R4. A couple is standing in a doorway. The man appears to be leaving; the woman, to be pleading with him. There is a poster with a semi-nude figure on it on the wall beside the door.

M5. A middle-aged woman is standing on the threshold of a half-opened door looking into a room.

R5. Here is a well-furnished living-room. There are no people in the room, but there is a woman looking into the room through a door she has half-opened.

M6BM. A short elderly woman stands with her back turned to a tall young man. The latter is looking downward with a perplexed expression.

R6BM. A well-dressed man, looking toward the floor, is standing behind a woman who has her back toward him. She is holding a handkerchief in both hands and looking in the direction of some drapes which could decorate a window. The man is holding his hat in both hands.

M6GF. A young woman sitting on the edge of a sofa looks back over her shoulder at an older man with a pipe in his mouth who seems to be addressing her.

R6GF. On a sofa is seated a woman who is looking over her shoulder at a man who is smoking a pipe. They appear to be talking.

M7BM. A gray-haired man is looking at a younger man who is sullenly staring into space.

R7BM. Two men are sitting beside each other. One of them is looking at his companion who is looking straight before him.

M7GF. An older woman is sitting on a sofa close beside a girl, speaking or reading to her. The girl, who holds a doll in her lap, is looking away.

R7GF. A woman is seated on a sofa near a young girl who has a doll on her lap. The woman seems to be reading or talking to the girl. The girl is looking away from the woman.

M8BM. An adolescent boy looks straight out of the picture. The barrel of a rifle is visible at one side, and in the background is the dim scene of a surgical operation, like a reverie-image.

R8BM. A teen-age boy is day-dreaming; a gun leans against the wall near him. He pictures a scene in which a man is lying on a cot; two more men are beside the cot, and one of them has a knife in his hand.

M8GF. A young woman sits with her chin in her hand looking off into space.

R8GF. A woman, chin on hands, is sitting in a chair; she is absorbed in thought.

M9BM. Four men in overalls are lying on the grass taking it easy.

R9BM. Four roughly-dressed men are lying on the ground in attitudes of relaxation. They form a close group.

M9GF. A young woman with a magazine and a purse in her hand looks from behind a tree at another young woman in a party dress running along a beach.

R9GF. A woman with a purse and a magazine in her hand is looking from behind a tree at another woman in party clothes, who is running along a beach.

M10. A young woman's head against a man's shoulder.

R10. A man and a woman are embracing each other, but they are not looking into each other's faces or kissing.

M11. A road skirting a deep chasm between high cliffs. On the road in the distance are obscure figures. Protruding from the rocky wall on one side is the long head and neck of a dragon.

R11. Everything here is rather vague but there seems to be a stream with rock-strewn banks and a rough bridge. There is a procession of figures which might be either animal or human or both. There is an indistinct, branching path on the far side of the stream. There is also a cave in the background from the mouth of which protrudes the head and long neck of a prehistoric animal.

M12M. A young man is lying on a couch with his eyes closed. Leaning over him is the gaunt form of an elderly man, his hand stretched out above the face of the reclining figure.

R12M. A man with closed eyes is lying on a couch. An older man is kneeling with one knee on the edge of the couch; he lets one hand hang in a relaxed position just above the face of the man on the couch.

M12F. The portrait of a young woman. A weird old woman with a shawl over her head is grimacing in the background.

R12F. A woman sits absorbed in her thoughts while a weird, old woman hovers just behind her.

M12BG. A row boat is drawn up on the bank of a woodland stream. There are no human figures in the picture.

R12BG. A row boat has been drawn up on the wooded bank of a stream. There are no people in this scene.

M13MF. A young man is standing with downcast head buried in his arm. Behind him is the figure of a woman lying in bed.

R13MF. This is a poor room which is furnished with a table holding a lamp and some unopened books, a chair, and a low cot. Lying on the cot with her face averted, is a woman who is nude to below the breasts. In the foreground, stands a man with his arm held across his face in such a way as to cover his eyes.

M13B. A little boy is sitting on the doorstep of a log cabin.

R13B. A little boy is sitting on the doorstep of a log cabin. He is leaning in such a way that his thumbs are pressed against his lips.

M13G. A little girl is climbing a winding flight of stairs.

R13G. A little girl is climbing a winding stair. Her left hand is resting on the stair rail. She is wearing a cap.

M14. The silhouette of a man (or woman) against a bright window. The rest of the picture is totally black.

R14. A figure is silhouetted in a long window. The figure may be either male or female.

M15. A gaunt man with clenched hands is standing among gravestones.

R15. This scene is set in a cemetery: there are irregular rows of tombstones. There is an old-looking figure in a knee-length, black coat. Its hands are clenched and pointed toward the ground.

M16. Blank Card.

R16. You are to imagine a picture; you are to describe it and to make up a story about it just as you have done for the others.

M17BM. A naked man is clinging to a rope. He is in the act of climbing up or down.

R17BM. A nude, muscular man is holding on to a rope with his hands. He could also use his bare feet to hold with, but the rope is hanging free between his feet. Although neither end of the rope can be seen, it is obviously hanging from some over-head point of attachment.

M17GF. A bridge over water. A female figure leans over the railing. In the background are tall buildings and small figures of men.

R17GF. A woman is leaning over a bridge railing. In the background, there are some buildings, and some men can be seen as they load a boat.

M18BM. A man is clutched from behind by three hands. The figures of his antagonists are invisible.

R18BM. This is a hatless man in an unbuttoned overcoat. His face is turned to look in the direction from which he has come, but his eyes are closed. Although there are no other people to be seen, he is being held by three hands.

M18GF. A woman has her hands squeezed around the throat of another woman whom she appears to be pushing backwards across the banister of a stairway.

R18GF. Two women are on a stairway. The hands of one are holding the neck or shoulder of the other who is leaning backward against the railing. It is not clear whether one is supporting or attacking the other.

M19. A weird picture of cloud formations overhanging a snow-covered cabin in the country.

R19. This is an abstract picture. Everything is blurred in outline as if with snow, and wind is indicated by jagged projections on several blurs; these projections stream in the same direction. There is a mound of white with two, lighted, squarish spots which might be windows or doors. There are two more lighted squares high above the first two. There are blobs and splotches of gray at the top of the picture.

M20. The dimly illumined figure of a man (or woman) in the dead of night leaning against a lamp post.

R20. A figure in a hat and long coat is leaning against a post with light shining down on it. The background is very obscure. There is an enclosure with low-growing plants behind the figure, and there are flecks of light which might be the lights of distant houses.

Rankin prefers a content analysis approach although she has used that of formal scoring occasionally. She has found her VTAT to be "an excellent projective technique."[4] Further, she has "observed no significant difference between the stories of blind persons and sighted persons even when the latter responses were made to the pictures instead of to the verbalization."[5] Her observations, however, are based on impression or inspection rather than on statistical evaluation.

Information about Rankin's VTAT has been briefly indicated previously.

D. Lebo's Use of Murray's TAT Descriptions

Independently, though at a later date, Lebo (1955) noted that the 10 TAT pictures which were most disliked because they contained affectively disturbing features were the sames ones whose descriptions contained disquieting elements. This observation suggested the possibility that the original descriptions of the TAT cards contained approximately the same stimulus value as did the cards themselves.

This notion received additional support in an investigation of Newbigging's (1955) concerning London college students who rated 10 TAT cards on a happiness-sad scale. When the ranks assigned by the London students were correlated with the identical cards from Lebo's study a rho of .92 resulted. This comparison suggested that in different investigations with subjects from different lands the TAT cards were responded to with similar affect: Newbigging's sad cards were Lebo's disliked cards.

---

[4] Rankin, S. K. Personal communication. May 1, 1958.

[5] Rankin, S. K. Personal communication. June 12, 1958.

The accuracy of the hypothesized relationship between the stimulus value of the verbal descriptions of TAT cards and the plates themselves was later examined by Lebo and Harrigan (1957). They compared protocols given to TAT cards for college girls with the responses given by the same subjects to a vocal presentation of Murray's descriptions.

To compare the value of such stories to traditional TAT protocols they used various objective measures which included word count, idea count, emotional tone of stories and outcome, response level, dynamic content, perceptual range, common themes, and normative data. All of these devices have been successfully employed in the traditional TAT literature. Scoring manuals (Lebo, 1958) defining each of the above techniques with references, illustrations of use, and with scored protocols embracing two sets of stimuli for adult females, can be obtained from the present writer.

The TAT and VoTAT protocols were found to be similar according to the measures employed. A later study by Lebo and Sherry (1959) contrasted ViTAT and VoTAT protocols of female college students using the scales and scoring methods listed above. A visual presentation of the descriptions yielded responses not unlike those given to a vocal presentation.

A study, now in process by Lebo and Naar, is examining the ViTAT and TAT responses of delinquent girls by means of transcriptions of tape recordings. Because of a high correlation between antisocial behavior, low-reading skill, and incapacity for self-expression (Salisbury, 1958) it will be interesting to observe whether the reading necessary for Murray's descriptions will reduce the similarities previously obtained for the TAT and VTAT responses of college girls. Should such a distinction not be the case it would suggest that the unchanged descriptions would have projective effectiveness even among poor readers. Preliminary examination of the data already collected from 10 institutionalized girls reveals no obvious differences.

A comparison of Murray's VTAT with blind and sighted subjects is being carried out by Maddox[6] using Lebo's scoring manual for methodology. The statistics are in process now.

The comparability of results from TAT, VoTAT, and ViTAT is, perhaps, not entirely unanticipated. Stone compared the responses of college students who listened to recorded sounds to the protocols of subjects who imagined the sounds. He reported that: "The use of imaginary (verbally described) sounds seems as productive, though less structured than the recorded stimuli" (Stone, 1955, p. 254). Similarly, Sargent (1956) has remarked that responses to the Insight Test were not markedly influenced by a vocal or visual manner of presentation.

---

[6] Maddox, J. K. Personal communication. April 16, 1959.

E. Conclusion

Both traditionally and empirically, then, a pictureless TAT administered visually or vocally by means of card descriptions would seem to be as diagnostically useful as the traditional TAT. Data obtained from some of the studies herein have already proven useful in answering questions raised about the diagnostic value of a pictureless TAT (Lebo, 1959). Further investigation if not the cautious use of VTAT's in clinical settings seems warranted.

References

1. Bellak, L. The Thematic Apperception Test in clinical use. *In* L. E. Abt and L. Bellak (*Eds.*), *Projective Psychology.* New York: Knopf, 1952. (Pp. 185-229.)
2. Lebo, D. Immediate affective reaction to TAT cards. *J. Clin. Psychol.*, 1955, **11**, 297-299.
3. ———. Scoring the verbal TAT, with illustrative protocols. Dittoed: Author, 1958.
4. ———. An empirical approach to problems concerning the diagnostic value of a pictureless TAT. *J. Proj. Tech.*, 1959, **23**, 107.
5. Lebo, D., & Bruce, R. S. Projective methods recommended for use with the blind. *J. of Psychol.*, 1960, **50**, 15-38.
6. Lebo, D., & Harrigan, M. Visual and verbal presentation of TAT stimuli. *J. Consult. Psychol.*, 1957, **21**, 339-342.
7. Lebo, D., & Sherry, P. J. Visual and vocal presentation of TAT descriptions. *J. Proj. Tech.*, 1959, **23**, 59-63.
8. Morgan, C. D., & Murray, H. A. A method for investigating fantansies. *Arch. Neur. & Psychiat.*, 1935, **34**, 289-306.
9. Murray, H. A. Thematic Apperception Test Manual. Cambridge: Harvard Univ. Press, 1943.
10. ———. Uses of the Thematic Apperception Test. *Am. J. Psychiat.*, 1951, **107**, 577-581.
11. Newbigging, P. L. Influence of a stimulus variable on stories told to certain TAT pictures. *Canad. J. Psychol.*, 1955, **9**, 195-206.
12. Salisbury, H. The Shook-Up Generation. New York: Harper, 1958.
13. Sargent, H. D. Insight Test prognosis in successful and unsuccessful rehabilitation of the blind. *J. Proj. Tech.*, 1956, **20**, 429-441.
14. Stone, D. R. Responses to imagined auditory stimuli as compared to recorded sounds. *J. Consult. Psychol.*, 1955, **19**, 254.

Child Guidance & Speech Correction Clinic
625 Ocean Street
Jacksonville 2, Florida

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 205-208.

COGNITIVE DISSONANCE AND YIELDING BEHAVIOR*

Department of Psychology, Montana State University

D. GORFEIN, T. KINDRICK, Q. LELAND, M. E. MCAVOY, AND J. BARROWS

A. INTRODUCTION

Festinger's theory of cognitive dissonance (3) has been applied to the understanding of behavior in some group influence experiments (Gilchrist, 4). Little work has been done however, in the manipulation of dissonance producing or dissonance decreasing variables in group influence situations of the form developed by Asch (1).

In examining the Asch group influence situation it would seem that one of the dissonance producing factors is the individual's problem as to whether to be true to his task (i.e., objectively reporting his perceptions) or to his need to be accepted by the group (i.e., yield to the group's standards).

In this experiment an attempt was made to increase the dissonance in the Asch situation by offering a reward for "getting all the answers correctly," (i.e., subjects were told that by getting all the judgments right they would receive a free ticket to a local movie theatre). If we did succeed in increasing the dissonance, we might hypothesize: (*a*) there would be less yielding by the experimental group (more subjects not yielding at all) due to reward motivation; (*b*) where yielding did occur in the reward group, it would be of non-compromise form, i.e., the subject would be committed to the group being objectively correct on those trials.

B. METHODS AND PROCEDURE

1. *Materials*

The stimulus cards used in this experiment were the same as those of Asch (1 p. 6).

2. *Procedure*

Following Blake and Brehm (2) a tape recorded version of the influence situation was employed in the following manner: (*a*) Upon arrival at the pre-arranged testing area (a wing of the psychology building) the subject was greeted by a pair of experimenters (all of the authors served in this capacity) with the question "Are you Mr. (Miss) Williams?" When the

---

* Received in the Editorial Office on June 9, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

subject denied this the experimenter then said "Oh then you must be Mr. ————" giving the correct name of the subject. He then took the subject into the testing room, instructing the other member of the team to wait for the fictitious Mr. Williams. (*b*) Upon being seated before a microphone and next to a loudspeaker, the subject was told that we were testing six subjects at a time. They were told that they were put into separate rooms so that they might all view the same card at the same time from the same angle. (*c*) Following this a voice on the speaker from a tape recording in an adjoining room made this announcement:

> This is to be a task involving the discrimination of the lengths of lines. Before each of you is a letter. This letter will designate the order in which you are to speak— Subject *A* goes first and so on down through letter *F*. Is this clear? Answer according to your letter now. Letter *A*? Female voice: "Does letter *A* mean that I always go first?" "Yes that's right." Letter *B*? "Yes" male. "Letter *C*?" "Yes" female. "*D*?" "Yes" female voice. "*E*?" "Yes" male voice. "Letter *F*?"

Subject *F*, the critical subject, invariably said yes. (*d*) The remainder of the instructions were the same as those of Asch (1 p. 4) with the exception that it was again repeated that all subjects would judge the same lines at the same time. Also half of the subjects (the reward group) were told that *all* subjects in their group who got all the matchings correctly would receive a free ticket to a local theatre. (*e*) Following this the 18 trials were run following Asch's procedure (1 p. 6) with each subject being called on in turn—the critical subject always last. (*f*) An interview followed the experimental portion. The questions reported by Asch (1 pp. 25-26) were asked.

3. *Subjects*

Twenty-two male students and 18 female students at Montana State University were divided into two equal groups (reward-control) of 20 students each.

C. Results

Hypothesis one, that there would be less yielding in the reward group was not confirmed ($X^2 = .40$ $P > .50$). There was no significant difference as to the number of individuals who yielded in each group (Table 1).

Hypothesis two, that where yielding did occur in the reward group it would more likely be of the non-compromise form (on those trials where compromises were possible[1]) was confirmed ($X^2 = 4.44$, $P < .05$) (Table 2).

---

[1] Compromise errors as defined by Asch are possible on Trials 3, 5, 6, 9, 11, and 12 (1 p. 17).

The data in Table 3 represents partial analysis of the interview material in an attempt to account for our failure to obtain as much yielding as that reported by Asch. There is a trend in the direction of those subjects, who reported that they thought the "others" might be looking at different cards, to be less likely to yield. Further testing among those who did mention this possibility and still yielded as compared to the other yielding subjects showed significantly more yielding on the part of those who did not mention this possibility (Mean 2.38, *SD* 2.45; in the mentioning group; Mean 5.30, *SD* 3.40, in the non-mentioning group; $t = 2.00$, $P < .05$).

TABLE 1
ERRORS MADE BY SUBJECTS

	Experimental (N 20)	Control (N 20)
Number making errors	8	10
Mean (all subjects)	1.75	1.85
Standard deviation	3.41	2.65
Mean per cent	14.60	15.40

TABLE 2
COMPROMISE ERRORS*

	Errors on compromise trials	Compromise errors
Experimental group	17	5
Control group	15	10

* $X^2 = 4.44$, $P < .05$.

TABLE 3
SUBJECTS HYPOTHESIZING DIFFERENT CARDS FROM OTHER GROUP MEMBERS*

	Yielding	Not-yielding
Hypothesizing	8	15
Not hypothesizing	10	7

* $X^2 = 2.28$, $P < .20$.

D. DISCUSSION

Failure to confirm our main hypothesis suggests the possibility that the magnitude of the reward offered was not great enough to increase the motivational dissonance in the reward group enough so that they might be more motivated to be correct than to maintain status with the group. This would suggest that work might be done, in terms of offering larger incentives, to further check this hypothesis. In fact, could we not measure, assuming that this hypothesis would be confirmed with larger incentives, the strength of group influence on people, by measuring the amount in terms of money it would take to get them to give objectively correct answers every time?

Failure to obtain as much yielding as Asch (1) in our situation can at least partially be accounted for by the availability to our subjects of an additional dissonance reducing device. Some of our subjects were able to rationalize their disagreements with the group by maintaining, despite instructions to the contrary, that the others might be responding to different cards. Analysis of the interview material did show that these subjects did yield significantly less often.

E. SUMMARY

Two hypotheses related to Festinger's theory of cognitive dissonance in terms of the effect of reward for correct judgments on a modified Asch group influence situation was tested. Partial support for these hypotheses was obtained. The effect of modification of the situation is also discussed in terms of dissonance reaction.

REFERENCES

1. ASCH, S. E. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol. Monog.*, 1956, **70**, No. 9.
2. BLAKE, R. R., & BREHM, J. The use of a tape recording to simulate a group atmosphere. *J. Abn. & Soc. Psychol.*, 1954, **49**, 311-313.
3. GILCHRIST, J. C. Social Psychology and group processes. In *Annual Review of Psychology*, 1959, **10**, Palo Alto: Annual Reviews, Inc.
4. FESTINGER, L. A Theory of Cognitive Dissonance. Evanston, Ill.: Row, Peterson, 1957.

Department of Psychology
Montana State University
Missoula, Montana

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 209-210.

THE EFFECTS OF DIFFERENTIAL POST-EXPOSURE ILLUMINATION ON THE DECAY OF A MOVEMENT AFTER-EFFECT*

Department of Psychology, Temple University

IRWIN M. SPIGEL

Exploration of a somewhat neglected parameter in the perception of movement after-effects appears potentially capable of providing a new approach to the understanding of neural involvements in these phenomena.

A recently completed pilot study has demonstrated the effects of post exposure darkness on the decay-rate of the after-effect. A systematic study of this influence, it is believed, will provide an instrument capable of more precise determination of whether a peripheral or central role predominates in the production of this form of apparent movement.

Specifically, it was found that an interval of darkness following exposure to a rotating spiral was in some way associated with a delay or inhibition of the decay of the usual after-effect of such rotation.

Ten male *S*s, all graduate students in psychology at Temple University, were employed for the study. Each *S* was given 10 trials, with a 30-second exposure to an 8-inch logarithmic spiral rotating at 200 rpm under conditions of constant illumination. As instructed, following cessation of rotation, *S* reported when he no longer perceived the after-effect. For each *S*, the mean duration of perceived after-effect was computed. A week later, each *S* was returned to the laboratory for a second series of 10 trials. At this time, following the 30-second rotation of the spiral, the laboratory remained in total darkness for a period of time equal to the mean duration of after-effect for the particular *S* on his first series of trials. The lights were then turned on, and as instructed prior to the session, *S* again reported when he no longer perceived any after-effect.

If the decay of the after-effect were independent of post-exposure illumination conditions, and related only to neural events in either the retina or brain, there should have been no statistically significant difference between the duration of the after-effect perceived by the *S* on his first series of trials, and that period of time represented by the darkness interval (*S*' mean after-effect) added to any additional after-effect perceived following re-lighting of

---

* Received in the Editorial Office on June 10, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the room in the second series. In short, the decay should have been approximately complete.

Instead, however, a significance test of the differences yielded a "*t*" of 3.7 (9df), with p less than .01. It appeared that the interval of darkness had in some way been accompanied by a delay or inhibition of the normal decay function.

In considering the implications of the findings, it occurred to the writer that the direct effect of light itself might be a necessary condition for the decay of the after-effect—or rather the cessation or inhibition of those neural events that account for the perception of the after-effect. It seemed conceivable that the neural involvements in after-effect phenomena depended to a large extent on continued stimulation by light energy; with the neural activity in question inhibited in the absence of this light energy. Should such be the case, it was further possible that some capacitor-type neural function was operative—whereby energy from stimulation builds up in nerve cells, and is dissipated (in after-effects) under continued activation under changed stimulus conditions. In the absence of light energy, the build-up may be prevented from discharging, and subject only to a determinable rate of "leakage."

In other words, the question may now be raised as to whether the presence or absence of illumination—or the presence or absence of interpolated contours—is chiefly responsible for the altered rate of decay. Preparations are now under way for a more systematic study of this question. The decay function will be studied under post-exposure conditions of darkness as well as of varying intensities of diffuse, unpatterned illumination.

Once established with some measure of confidence, the significant parameters in the decay function will be utilized under several interocular conditions in an effort to determine the relative contributions of central and peripheral influences in the production of the after-effect.

Box 416 (Room 822)
Temple University
Philadelphia 22, Pennsylvania

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 211-225.

ALTERNATION LEARNING AND MENTAL PATHOLOGY: TEST PROCEDURES AND FINDINGS*

Veterans Administration Center, Hot Springs, South Dakota

MILTON B. JENSEN[1,2]

A. INTRODUCTION

The assessment of learning via alternation procedures has an impressive history through extensive clinical application is relatively recent. Where or when alternating presentation of learning tasks appeared in the laboratory (or in pedagogy, for that matter) is not known. Some of the T-Mazes that were almost standard equipment in the psychology laboratories of the 1920s presented choices in a single position (always to the right or always to the left). Some used single alternation presentation (left, right, left, right, or the reverse). Hunter introduced the term formally in 1928 in "The behavior of racoons in a double alternation temporal maze." In 1936 J. McV. Hunt reported what probably is the earliest formal clinical assessment of disordered persons via the double alternation. He presented *S* with five playing cards face down on a table. From these *S* was to find the key card, e.g., the ace of spades, the position of which followed the sequence left, left, right, right in consecutive presentations. Two consecutive errorless completions of the sequence constituted "success." Most of the studies on alternation learning since the reviews of Munn in 1950 and Stone in 1951 may be traced directly to the influence of Pascal at the University of Tennessee who studied for the Ph.D. with Hunter. The alternation studies of Pascal and Jenkins (1959) stimulated those of Jensen (1957, 1958a,b). Pascal and Jenkins developed a bank of five boxes with which varied alternation and delayed reaction tasks could be presented. They called this the Hunter-Pascal Concept Formation Test (1959). Jensen developed his Alternation Board described in *The American Journal of Psychology* (1958a) and manufactured and distributed by the Lafayette Instrument Company (1958,b). The present study is concerned with the development of alternation testing procedures (playing cards to boxes to the Jensen Alternation Board) and with data for the patients with whom psy-

---

* Received in the Editorial Office on June 13, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Volunteer workers, Mrs. W. D. Bolls and Mrs. J. L. Goldman, administered most of the tests used in this study. Their help is gratefully acknowledged.

[2] I am deeply indebted to Dr. Carol H. Ammons, Editor, *Psychological Reports* for many suggestions leading to preparation of this paper in its final form.

chologists worked most while the writer was stationed at the VA NP Hospital, Salisbury, North Carolina.

B. Recent Clinical Applications

In 1948 Pascal and Zax reported on use of the double alternation test with cerebral palsied children. Hodges reported age norms for the double alternation on 260 boys in the Oak Ridge, Tennessee schools, ages six to 12, inclusive, with socio-economic and intellectual ability "controlled" (1954, 1956). In 1957 Leslie completed a doctoral study at the University of Tennessee using the double alternation with hospitalized psychotics. In 1957 Jensen reported preliminary findings on the performance of NP hospitalized veterans on alternation tasks of several degrees of difficulty. Sanders and Weinman reported alternation test data from the Philadelphia State Hospital in 1959. Students of Pascal and Jenkins at the University of Tennessee completed several studies during this period: Collins in 1957, Davis in 1958, Kaiman in 1958, Larsen in 1958 and Leslie in 1956 and 1957. Pascal and Jenkins reported on the Hunter-Pascal Concept Formation Test (alternation and delayed reaction) giving test procedures (1959,a) and normative data for humans, ages two months to 22 years (1959,b). At the end of the first paragraph of the latter report they say: "Performance on both problems has been shown to be related to position on the phylogenetic scale, using animals from fish to humans, and ontogenetically, to maturation level in human subjects."

C. Methods of Presentation—The Alternation Board

1. Presentation via playing cards (McV. Hunt's method, 1936) was discontinued in 1955 when it was found that several *S*s made near-perfect scores by marking the cards to be located.

2. Presentation via a gang of five boxes (Pascal-Jenkins method, 1959) with doors facing *S* was unsatisfactory with our NP patients. It was impossible for *E* always to place the object to be sought without giving cues and at the same time maintain a non-threatening test situation for our mental patients.

3. Development of the Alternation Board (Jensen, 1957; 1958a; 1958b) provided a method of presentation which, in our experience, precludes cheating by *S* and the giving of cues by *E*. It is interesting and very seldom threatening to mental patients—about 5 per 1,000 in our experience. It consists of a bank of five lights and five corresponding switches mounted on a base (10 in. by 14 in.). A switch will turn on its corresponding light when acti-

vated through a multi-position control switch which is concealed from *S* by a screen (5-in.) sloped (130°) away from *S*. By changing the position of the control switch, *E* may in one complete clockwise revolution present four alternative tasks: single position (the same light each time), single alternation, double alternation, and 5-position alternation. Power is supplied by one flashlight battery. The learning tasks may be altered in pattern and complexity by changing the wiring on the control switch. Alternate forms of the task are obtained by turning the control switch counterclockwise.

D. TESTING AND SCORING PROCEDURES

1. *Instructions to S*

Idealistically, we would have preferred to avoid written and spoken instructions. Since this was impracticable, we compromised with: (*a*) A simple statement as to the nature of the task at the beginning of the testing. (*b*) Warning of the beginning of each sequence by the words, "Let's try again." (*c*) Warning of the beginning of each new alternation task by, "Now, the problem is different" (Jensen, 1959).

2. *Number of Trials*

The number of trials per alternation task was cut from 20—the number used by Hodges (1954)—to 10. There was a correlation of .78 between the first 10 and second 10 of 20 trials of 173 *S*s and none of the 173 succeeded on the second 10 of 20 trials after failure on the first 10 of the double, 5-position or 7-position alternations.

3. *Success*

Success on an alternation task was arbitrarily defined as two consecutive errorless trials of a sequence, providing *S* could, in any way, demonstrate knowledge of the sequence. If *E* was not satisfied that *S* knew the sequence, the trials were continued until *E* was satisfied or until the end of 10 trials, whichever came sooner.

4. *Scoring for Success and Failure*

We experimented with several methods of scoring. Of these, our extension of Hodges' method (1954) was the most discriminatory and predictive. Hodges demonstrated a significant mental-age difference between boys who passed and boys who failed the double alternation with the same number of errors. Consideration of success and failure seemed justified with our NP patients also. With 241 *S*s there was an overlap of error scores of those who succeeded and those who failed as follows: single position, 39 per cent; single alternation, 52 per cent; double alternation, 47 per cent; 5-position alternation, 38 per cent. Adjustment for success and failure by Hodges' method (1954) reduced the overlap of these distributions of scores to 12, 14, 16, and 13 per cent. If, for instance, *S* made 28 errors in 10 trials on the double alternation, using Hodges' procedures (1954), he scores MA 9.7 years, if he learned the sequence, and MA 8.6 years if he did not learn it. With success he would score at the 55 percentile of our patients; with failure, at the 44 percentile. The advantage of allowing for success and failure is further demonstrated by standard errors of individual scores. These are significantly larger with failure than with success on each of our four series and larger for total errors regardless of success or failure, than for total score in which success-failure is considered. The standard error of a single total score as weighted in Table 1 is 3.7 percentile points. For a total error score regardless of success or failure it is 6.00 percentile points.

5. *Fixation of First Position*

Hodges provided a graduated penalty for errors after fixation of the first position of the sequence of the double alternation (1954). We did not use this correction in scoring the performance of our patients since, in most instances, it amounted to not more than one or two percentile points.

6. *Scoring for Time*

Time to completion of the test was so much a function of the speed of *E* that it was abandoned as a scoring device. We think the speed of *S* would be very significant were it measured separate from the speed of *E*. We have not found it practicable to do so despite a Pearson *r* of .78 between time to completion and mental age on the double alternation (using records of 173 examinations by two *E*s). The time actually used by different *E*s in giving instructions, recording, and resetting the switches varied as much as 50 per cent. Our fastest examination with all four series was completed in five minutes. The slowest took 70 minutes. The average for 173 was 13.3 minutes.

7. *Scoring Success after Failure*

If *S* passed a more difficult series after failure on an easier one, we credited him with the average of his percentile scores on the series subsequent to the early failure. If, for instance, *S* failed the single alternation (*A*) at the 20 percentile, passed the double alternation (*B*) at the 50 percentile, and failed the 5-position task (*C*) at the 46 percentile, we credited him with percentile

48 (the average of *B* and *C*) on *A* and used the weight for that percentile in arriving at his total percentile score (See Table 1).

8. *Scoring Inability and Refusal*

Arbitrarily, we scored those who were too confused to take the test and those who refused to do so at the 0.0 percentile. The predictive value of these behaviors is discussed later (p. 221).

9. *Scoring Repetitive Errors*

We found no satisfactory method of penalizing repetitive errors in our scoring scheme. We had similar difficulty scoring the performance of the *S* who, after repeated success with a response relinquished it in favor of one he had previously demonstrated to be unsuccessful.

10. *Uniformity of Success and Failure*

Thirty-eight of 231 (16.5%) learned the sequence on all series *A, B,* and *C*. Seventy-five of 231 (32.5%) failed to learn the sequence on either of the three series, *A, B, C*. We found no method of scoring these behaviors, though we think that they merit special attention.

11. *Alternate Forms*

In Form I of our series the correct choices for the five series, numbering 1, 2, 3, 4, 5 from *S*'s left are: Single Position, No. 2; Single Alternation, Nos. 3 and 4; Double Alternation, Nos. 1, 1, 5, 5; 5-Position Alternation, Nos. 2, 5, 2, 4, 2. Our alternate form (II) uses No. 4 from *S*'s left for the Single Position and the reverse of Form I of each of the other series. Presentation of Form I is accomplished by clockwise turning of the control switch of the Alternation Board; Form II by counterclockwise rotation. Means and standard deviations of errors on the two forms for two randomly selected groups of 50 each did not differ significantly for any series (*CR*s all < p,. 12). The groups were randomized by alternating the forms applied to 100 consecutively examined patients.

E. INTERNAL EVIDENCES OF MERIT

1. *Range of Difficulty*

Of 173 patients, the per cents learning the sequence for each of the series were: single position, 75.6; single alternation, 55.8; double alternation, 44.2; 5-position alternation, 21.0; 7-position alternation, 21.0. The 7-position alternation was eliminated from the series. The single position task was retained as a "warm-up" exercise. The other tasks were kept as forming a satisfactory series of increasing difficulty.

2. *Reliability*

Coefficients via the Spearman-Brown formula (errors on odd vs. errors on even trials) for 100 records selected without known bias were: single position, .61; single alternation, .97; double alternation, .96; 5-position alternation, .94. Reliability for the five series combined was estimated at .985 by Kelley's adaptation (1923, Formula 148) of Spearman's method for securing the reliability of weighted averages. Without the single position series total reliability by this procedure was .983.

F. NORMS FOR OUR MENTAL PATIENTS

From a population of 263, 22 were eliminated because they would not or could not take the test and 10 because they passed a more difficult series following failure on an easier one. Scores on the remaining 231 were cast into frequency tables according to number of errors for success and failure on each of the series. These distributions were treated by Cureton's 1956 7-point smoothing formula to preserve parabolic and cubic trends while smoothing out irregularities of higher orders. Percentile scores were determined for Hodges' mental ages on the double alternation. For these there was a six per cent overlap in mental age between the highest error score of those who passed and the lowest error score of those who failed. This same overlap was allowed in deriving percentile values for the other series and overlapping error scores were distributed evenly throughout the overlapping ranges on all the series.

Pearsonian intercorrelations between the percentile scores of the 231 *S*s on the single position, single alternation, double alternation, and 5-position alternation series, in order were .212, .224, .349, .535, .708, and .780.

Total percentile scores were derived by weighting the percentiles of the individual series, excluding the single position task, by Kelley's method (1923, Formula 148). They were 1.1916, 1.0604, and .5478 respectively for the single, double, and 5-position alternation tasks. The per cent that each of these weights was of their sum (2.7998) was used as a multiplier at each percentile level. These products produced the weights opposite the error scores of *S* on the three series (Table 1). Added together they provide a total percentile score. Normative data for our mental patients are given in Table 1.

TABLE 1

Alternation Test Score Norms for NP Hospitalized Veterans

Percentiles are given at the left opposite errors under X+ (for success) and X— (for failure) for each series. a = single position; A = single alternation; B = double alternation; and C = 5-position alternation. Total percentile score is the sum of the weights opposite the number of errors on series A, B and C. MA = mental age by Hodges' formula (1954). CW is the approximate percentile for 45 undergraduate college women. Means and standard deviations for the group are in errors. The standard deviations of individual scores are in percentile points.

	a Single position		A Single alternation				B Double alternation					C 5-Position series		
%ile	X+	X—	X+	X—	Wt.	CW*	X+	X—	Wt.	MA	CW*	X+	X—	Wt.
100	1		1		43		2		38	11.5		6		19
95	2		3		41	95	5		36	11.2		18		18
90	3		5		39		8		34	10.9	50	27		17
85	4		6		37	80	11		32	10.7		39	24	16
80	4		7		34		14		30	10.5	30	54	33	16
75	5		8		32	50	15		29	10.4	10		41	14
70	5		9		30		17		27	10.3	5		45	13
65	5		10		28	30	20		25	10.1			49	12
60	6		12		26	20	22	12	23	10.0	2		53	11
55	7		14		24		28	15	21	9.7			58	10
50	8		16		22	15	40	20	19	9.2			63	9
45	8		18		19			27	17	8.7			68	9
40	9		21		17	10		33	15	8.2			73	8
35	10		26	16	15			39	13	7.9			79	7
30	11		34	22	13			48	11	7.4			84	6
25	12			26	11			63	10	6.7			92	4
20	14	7		29	9			74	8	6.2			96	3
15	18	10		32	7			81	6	5.9			101	2
10	23	15		36	4			90	4	5.6			105	2
5		23		46	2			100	2	5.3			111	1
0		31		52	0			126	0	4.9			138	0
%+			87		68					47			17	Total
M (errors)			47		48					49			49	48
SD (errors)			27		28					23			30	28
**SD (score)			16		5.0					4.5			7.0	3.7

* Approximate %ile for 45 undergraduate college women, University of Tennessee.

** Standard error of a single percentile score.

G. Correlates

Most of the correlations given in Table 2 are for the 159 patients examined in 1956 and 1957 and followed to December 1, 1959. Long-term refers to 32 long-term closed ward male patients. The influence of these on data for the total population is reflected in the correlations.

TABLE 2
Correlations for Patients Followed 31 Months or More After Testing

	Total (N = 159)	Less Long-Term (N = 127)	Long-Term (N = 32)
With alternation test scores:			
Per cent of time spent out of the hospital	.50	.43	.28
Per cent of time spent on closed wards	—.51	—.38	—.63
Rating by examiner (N = 173)	.83		
Wechsler-Bellevue Verbal IQ (N = 141)	.13		
Porteus IQ (N = 64)	.07		
Otis IQ (42 college women)	.08		
Sex (N = 173)	.02		
School grades completed	—.13		
With % of time spent out of the hospital:			
Married vs. single	.59	.21	.71
Ever married vs. never married	.48	.25	.27
Single vs. divorced or separated	.39	.00	.34
Married vs. divorced or separated	.29	.07	.88
Baptist vs. not-Baptist	.23	.00	.34
Sex	.22		
Race (white-other)	.19		.32

H. Discussion of Correlates

1. *Duration of Hospitalization*

Our correlation of .50 between alternation score and per cent of time spent out of the hospital is about as high as is reasonable. A multiplicity of factors operate, and not necessarily concomitantly, to keep the mental patient in or out of the hospital. Some of these are: Family solidarity, if any. Motivation of the patient. Staff and administrative policy, conscious and unconscious. Nowhere to go on leaving the hospital and lack of funds. Unfavorable attitudes towards the ex-mental patient. And, location of a patient on a ward with those who have not responded to treatment.

Because of these, and other factors, length of hospitalization cannot be predicated entirely upon degree of mental illness. We found, for instance, a near-perfect correlation between average alternation scores on various types of wards in the hospital and severity of mental pathology on those wards as seen by our professional staff. But, the correlation between average alternation score for those treatment areas and per cent of time the patients had spent out of the hospital in more than 31 months after testing approximated zero.

2. *Alternation Scores and Intelligence*

We have found no consistently significant relationship between alternation scores of our patients and their scores on any of a wide variety of intelligence and other tests. With our patients, the alternation test is not a conventional test of intelligence.

Reasonably direct comparisons were possible between the scores of 45 undergraduate college women at the University of Tennessee and our patients on the single and double alternation tasks. On the single alternation only 27.2 per cent of our patients scored as high as the average of the college women; on the double alternation, only 24.3 per cent. The average Otis IQ of the college women was 113.9 (SD, 8.8). The average Wechsler-Bellevue Verbal IQ of our group was 97.7 (SD, 15.0). An attempt was made to equate the scores of our patients with those of the college women to the extent that the two groups overlapped. These approximations are incorporated into Table 1.

3. *Intellectual Impairment*

We think that about half of our patients are severely impaired in learning ability. In arriving at this conclusion we compared their alternation scores with present scores on verbal and non-verbal intelligence tests and with premorbid achievements. We considered their academic, military, and occupational histories. Mostly we used the Wechsler and Porteus intelligence tests. We think that our patients, on the average, have lost the equivalent of about 20 IQ points in learning ability because of mental illness. We can see no other explanation for the learning deficiencies of some of our patients who: (*a*) completed 12 or 13 grades of school with better than average grades, (*b*) served honorably as officers or non-commissioned officers throughout WW II, and (*c*) scored above average on both the WAIS and the Porteus, yet, (*d*) failed the double alternation task with an error score equal to that of the average boy of age 6½ years. This is not an unusual type of performance for our patients.

Many of them scored as low as mental defectives on the alternation test. However, the alternation performance of the low scoring mental patient and of the mental defective are radically different. The mental defective seems to learn—but not much—in an orderly sequence and to retain what he learns. The mental patient with comparable low alternation scores makes many repetitive and irrational errors and often rejects choices he has previously found appropriate.

4. *Type of Mental Illness and Alternation Score*

Mean alternation scores of two diagnostic groups of patients differed at beyond the .001 level of confidence. They were: (a) 118 schizophrenics and manic depressives; (*b*) 35 alcoholics, anxieties, and personality and character disorders. Their average percentile scores were 42 and 59, respectively. Alternation score did not separate the schizophrenics from the manic depressives nor any of the sub-groups in (*b*).

5. *Influence of Long-Term Closed Ward Patients on Relationships*

When the long-term closed ward male patients were excluded from our total population, most of the classical correlates of duration of NP hospitalization were not significantly related (Table 2). Married vs. single and ever-married vs. never-married were not more predictive than sex, race, or Baptist vs. not-Baptist. The effect of the long-term closed ward patients on our data is shown by Table 3. It is most apparent when we exclude from the long-term closed ward roster those who scored percentile 50 or higher on the alternation test. Among our long-term closed ward patients there is a hard core. As far as we can tell, they have been schizophrenic from childhood or adolescence. Since they have had little opportunity and/or drive for many of the activities of young adults—marriage for instance—they distort the data relative to those activities for any group of which they are a part.

TABLE 3
EFFECTS OF LONG-TERM CLOSED WARD PATIENTS ON GROUP DATA

	N	Av. total alt. %ile	% of time out of hosp.*	% of pts. in hosp.**
All tested in 1956-1957	159	41.6	48	36
Long-term closed ward	32	33.9	12	75
Total less long-term closed ward	127	44.9	57	32
Long-term scoring above 50%ile	9	64.0	33	33
Long-term scoring below 51%ile	23	22.0	5	87

* From time of testing to December 1, 1959.
** As of December 1, 1959.

We were prepared for the wide range of performance on most of the wards but surprised that nine of the 32 long-term closed ward male patients scored above the 50 percentile of the entire group. It is striking that these nine (15.6%) who scored above the 50 percentile on the alternation test accounted for 70.3 per cent of the time spent out of the hospital by their entire group of 32 during the more than 31 months following examination.

I. CHARACTERISTICS OF EXTREME PATHOLOGY

1. *Refusal to Take the Test and Inability To Do So*

Of the 159 patients examined in 1956 and 1957, 11 refused to take the test and four were too confused to do so. None of those who refused to take the test spent any time out of the hospital from time of refusal to December 1, 1959—approximately 31 months. Only one of them spent any time off a closed ward during that period. The four who could not take the test spent 32 percent of their time out of the hospital during the same period and 16 per cent of it on open wards. In contrast, the group, exclusive of these 15, spent 51 per cent of their time out of the hospital during those 31 months.

2. *Many Errors*

Clinically, those who made many errors were strikingly different from the generality of patients and from those who made few errors. Fifteen patients who failed to learn the double alternation and averaged 78.3 errors (MA, 6.1 yrs.) were compared with 15 who learned it with an average of only 10.3 errors (MA, 10.7 yrs.). Histories and data from physicians, nurses, psychologists, social workers, and others agreed that these groups were very different in severity of pathology and not different in socio-economic background, amount of schooling, or premorbid vocational achievement. But, the impairment of several was not so apparent in ward behaviors. Only trained observers noted such things as "difficulty concentrating, short memory span," etc. The low-error group was made up exclusively of those diagnosed as having anxiety reactions, personality disorders, and depressive reactions. Those who made the most errors were all long-term schizophrenics with acute symptoms.

3. *Repetitive Errors*

Severe pathology was characterized by many errors on the alternation test and by repetition of the same errors, time after time—or trial after trial. Some pushed the same switch at the beginning of all trials despite the fact that it never once lighted the lamp. Some never learned the first switch to push in a series despite the fact that it was always the first switch to activate a light. Examiners said of them: "Just pushed switches," or "Repeated the same error over and over again." It was as though they projected the purposeless, non-judgmental, perseverative behaviors of their entire pathological histories into this single test situation. The record sheets of the non-learners were filled with numbers representing practically all the errors possible, and often in a random or strictly repetitive order. On the double alternation, 49 of 111 (44%) repeated their first errors an average of 4.59 times. The average number of such repetitions for the entire group was 2.07. For the long-term closed ward patients it was 4.71. (And they rejected their first successful responses on 61.2 per cent of their later choices! Two of them responded with 5—4—3—2—1 on all 10 trials on the sequence 1—1—5—5.) Even one repetition of the first error on an alternation task probably is significant. Fifty-six per cent of our patients never repeated their first errors on the double alternation. By chance alone, five consecutive repetitions of the initial error on an alternation task should occur only about once in 3,000 times. Five of our long-term closed ward patients repeated their initial errors every time they had opportunity to do so the double alternation—nine times in 10 trials.

4. *Success after Failure*

Ten of 241 patients learned a more difficult alternation sequence after failure on an easier one. We have found no satisfactory explanation for this phenomenon. It is about as common to non-mental-hospital groups as to our patients. In deriving total scores for such cases we average their percentile scores on the more difficult series and credit them with this on the easier series that was failed. Thus, with failure on *A* at the 10 percentile, success on *B* at the 55 percentile, and failure on *C* at the 65 percentile, the score for *A* is the average of 55 and 65 or 60.

J. ILLUSTRATIVE CASES

1. *Bright, Severe Impairment*

Male attorney. Age 63. No organicity. Diagnosed Manic Depressive Reaction, Manic Type. Successful professional, family and civic relationships most of life to 1955 except for periods of acute mental illness. Reported 10 hospitalizations from 1920 to 1955. From testing 5/22/56 to 12/1/59 (43 months): Two readmissions. Trial Visit, nine months. Discharge status, 24 months. On closed wards seven months. On open wards two months. Wechsler-Bellevue Verbal IQ, 120. Jensen Alternation Scores: Passed single alternation at 25 percentile with 31 errors. Failed double alternation at 8 percentile with 94 errors (Age equivalent, 5.2 yrs.). Failed 5-position alternation at 5 percentile with 111 errors.

2. *Bright, No Impairment*

Male bookkeeper. Age 34. Ten grades schooling. Diagnosed Anxiety Reaction. Guilt ridden over killing prisoners during WW II. Difficulty with

In-Laws. Good work history except for periods of incapacitation from excessive drinking. From testing 5/15/56 to 12/1/59 (42½ months): Hospitalized on two occasions (drunk on admission) for a total of 37 days. Wechsler-Bellevue Verbal IQ, 126. Jensen Alternation Scores: Passed single alternation at 85 percentile with six errors. Passed double alternation at 72 percentile with 15 errors. Passed 5-position alternation at 97 percentile with 19 errors.

3. *Dull, No Impairment*

Unemployed male laborer. Age 34. Five grades schooling. Diagnosed Chronic Brain Syndrome with Convulsive Disorder. From testing 4/24/56 to 12/1/59 (43 months): On Trial Visit 15 months. On discharge status one month. On closed wards 21 months. On open wards six months. Wechsler-Bellevue IQ, 63 (Verbal 75; Performance 55). Jensen Alternation scores: Passed single alternation at 95 percentile with three errors. Failed double alternation at 25 percentile with 62 errors (MA, 6.7 yrs.). Failed 5-position alternation at 55 percentile with 57 errors.

4. *Educational Deficit, Not Dull, Not Impaired*

Male on-the-job auto mechanic trainee. Age 25. Five grades schooling. Diagnosed Convulsive Disorder Grand Mal. From testing 1/31/56 to 12/1/59 (46 months): One readmission. On discharge status 43 months. Wechsler-Bellevue Verbal IQ, 63. Jensen Alternation Scores: Passed single alternation at 75 percentile with eight errors. Passed double alternation at 85 percentile with 11 errors (MA, 10.7 yrs.). Passed 5-position alternation at 99 percentile with four errors.

K. SUMMARY

This report summarizes about five years of work with alternation procedures in a Veterans Administration NP hospital. History, presentation, and methods of scoring are discussed and a table of norms is given.

In the opinion of the writer, the alternation test provides a unique method of measuring intellectual impairment. Many mental patients with high school and college training who score well on memory-type tests of intelligence learn the alternation tasks no better than does the average five- or six-year-old child. We think that the learning ability of the average patient at this hospital is impaired an equivalent of about 20 IQ points as compared with premorbid level. Our data lead to the not-hopeful conclusion that, in severe cases this deterioration is not reversible. We think that post-morbid adjustment, if made, will have to be on a level commensurate with the impaired learning ability and not relative to premorbid achievement. We see little prospect of rehabilitating the long-term mental patient (even though, premorbidly, he was a successful business or professional man) if he learns alternation procedures at a very low level.

REFERENCES

1. COLLINS, M. P. A study of the development of concept formation as related to the maturation of place response, delayed reaction, and single alternation in a selected group of children ages thirty-four to sixty months. M.A. thesis, Univ. Tenn., 1957.
2. DAVIS, H. C. A psychological study of chronic alcoholics. Ph.D. dissertation, Univ. Tenn., 1958.
3. CURETON, E. E. Seven-point smoothing formula. Unpublished ms., Univ. Tenn., 1956.
4. HODGES, A. A developmental study of symbolic behavior. *Child. Devel.*, 1954, **25**, 277-280.
5. ———. Double alternation: A measure of intelligence. *J. Consult. Psychol.*, 1956, **20**, 59-62.
6. HUNT, J. McV. Psychological experiments with disordered persons. *Psychol. Bull.*, 1936, **33**, 1-58.
7. HUNTER, W. S. The behavior of raccoons in a double alternation temporal maze. *J. Genet. Psychol.*, 1928, **35**, 374-388.
8. JENSEN, M. B. Learning ability of N. P. hospital patients as measured by modifications of the Hunter double alternation procedures. *Amer. Psychol.*, 1957, **12**, 366. (Abstr.)
9. ———. A light-switch alternation apparatus. *Amer. J. Psychol.*, 1958, **17**, 441-442. (a)
10. ———. The Jensen Alternation Board. Lafayette, Indiana: Lafayette Instrument Company, 1958. (b)
11. ———. Manual for the Jensen Alternation Board (Revised). Lafayette, Indiana: Lafayette Instrument Company, 1959. (Mimeographed)
12. KAIMAN, B. D. A developmental study of concept formation in children. Ph.D. dissertation, Univ. Tenn., 1958.
13. KELLEY, T. L. Statistical Method. New York: Macmillan, 1923.
14. LARSEN, E. R. A developmental study of performance on the Hunter-Pascal Concept Formation Test. Ph.D. dissertation, Univ. Tenn., 1958.
15. LESLIE, D. L. A methodological study of the Hunter-Pascal Concept Formation Test in a group of nursing students. Unpub. ms. Univ. Tenn., 1956.
16. ———. The relationship between performance on the Hunter-Pascal Concept Formation Test and the neurotic-psychotic dichotomy. Ph.D dissertation, Univ. Tenn., 1957.
17. MUNN, N. L. Handbook of Psychological Research on the Rat. New York: Houghton-Mifflin, 1950.
18. PASCAL, G. R. Behavioral Change in the Clinic—A Systematic Approach. New York: Grune & Stratton, 1959.
19. PASCAL, G. R., & JENKINS, W. O. The Hunter-Pascal Concept Formation Test: Directions for administering and scoring. Univ. Tenn., 1959. (Mimeographed)
20. ———. The Hunter-Pascal Concept Formation Test: An experimental approach to the measurement of cortical capacity. *J. Clin. Psychol.*, 1959, **15**, 159-164.

21. PASCAL, G. R., & ZAX, M. Double alternation performance as a measure of educability in cerebral palsied children. *Amer. J. Ment. Defic.*, 1955, **59**, 658-665.
22. SANDERS, R., & WEINMAN, B. Psychosocial treatment of the chronic mental patient: A rationale. *Amer. Psychol.*, 1959, **14**, 289. (Abstr.)
23. STONE, C. P. (*Ed.*) Comparative Psychology. New York: Prentice-Hall, 1951, (3rd Ed.).

Veterans Administration Center
Hot Springs, South Dakota

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 227-251.

THE EFFECT OF TRAUMATIC EXPERIENCES ON THE GROWTH AND BEHAVIOR OF THE RAT*[1]

Laboratory for Behavioral Research, Medical Center, University of Arkansas

CHARLES R. GALBRECHT, ROSCOE A. DYKMAN, AND JOHN E. PETERS[2]

A. PROBLEM

The present study developed from some accidental results obtained with one experimental dog in a previous study (Dykman and Gantt, 1960). During an early preconditioning period, three electrical stimuli of a very high intensity were applied to the animal's left forepaw. As a result of circuit difficulties, the shocks were higher in intensity than was intended. During the first two days following the shocks, no signs of disturbance other than an accelerated heart rate (HR) were observed. However, between the second and the 24th day, the following pathological responses appeared in the absence of further traumatic stimulation: pronounced tachycardia, trembling, motor restlessness, aperiodic struggling, cowering at the sight of the experimenter, urination, defecation, vomiting, and penile erections. Various procedures were tried in an attempt to extinguish these responses; e.g., special care and food, multiple daily sessions, and mild electrical stimuli. Only the mild stimuli were effective. After two days of mild stimulation, the above mentioned signs of emotional disturbance largely disappeared, and the animal was sufficiently recovered within a week so that his conditioning routine could be continued. The process of developing pathology in the absence of further traumatic stimulation has been termed autokinesis by Gantt (1953).

Our original intention was to repeat this work with dogs of different species. It was decided, however, that it might be best first to try a procedure similar to that described with smaller and less expensive animals than dogs, and rats were chosen as *S*s. Investigation of the literature revealed few studies

---

* Received in the Editorial Office on June 24, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by the National Institutes of Mental Health, Public Health Service, under Research Grant No. M-1091.

[2] The writers wish to acknowledge the technical assistance of Mrs. Margaret Townsend and Mr. Gene Townsend (junior medical student). We also wish to thank Dr. W. K. Jordan, Professor of Neurology, who provided us with laboratory facilities. All of the above mentioned people were associated with the University of Arkansas at the time of the study.

in which the behavior of rats subjected to traumatic stimulation was followed over long periods of time, and there appeared to be no longitudinal studies employing a large number of indices of disturbance.

The specific purposes of the present study were as follows: (*a*) to study the effect of differing numbers of electrical stimuli (intensity constant) upon various selected behaviors, upon autonomic function, and upon the growth of the young albino rat; (*b*) to compare certain behaviors and autonomic function in terms of the duration of disturbance; and (*c*) to study the "therapeutic" role of mild electrical stimuli in alleviating pathological symptoms. The general aim was the development of a standardized procedure for producing pathological changes in the rat that might be used in studying the effects of various types of "therapies" and intracranial stimulation.

B. METHOD

1. *Subjects*

Thirty albino rats of the Sprague-Dawley strain, 15 males and 15 females, were used. The rats weighed 200 to 250 gms.; the males were two months old and the females four months old upon delivery. The experiment was begun seven days after delivery.

The 15 *S*s of each sex were assigned at random to three experimental groups with five rats in each group. This procedure provided for the establishment of two control, two one-shock per day, and two three-shock per day groups. These groups will be designated *C*, *L*, and *H*, respectively, with the letter *M* or *F* denoting male or female. For example, *FC* will designate the female control group, *MH* the male three-shock per day group, *L* the *ML* and *FL* groups, and *H* the *MH* and *FH* groups.

The six groups of *S*s were housed in separate cages and allowed to drink *ad libitum*. At 5:00 p. m. each day 200 gms. of Purina Laboratory Chow were placed in each of the three male cages and 185 gms. in each of the three female cages. Unfortunately, there were no provisions made for temperature or humidity control in the animal colony in which the *S*s were maintained.

2. *Apparatus*

The major piece of apparatus was an observation box made of wood and painted flat gray. The box was 16 in. long, 14.5 in. wide, and 12 in. high. Except for a small triangular shaped door, the ceiling was covered with ½ in. wire mesh. The grid was comprised of copper strips ⅛ in. wide and spaced ½ in. apart (the copper strips were embedded in a plastic insulating material). Four fixed metal posts, each 4 in. high, located in the corners of the box, supported the grid. Two desk lamps placed above the mesh evenly illuminated the interior of the box. A variac with an isolating transformer arrangement delivered the electrical stimuli, which shall henceforth be designated as unconditional stimuli (US).

Motor movements were recorded by a phonograph cartridge fastened to the underside of the grid. The cartridge was connected to a Grass EEG and was sufficiently sensitive to register all but the smallest head-turning movements. Heart rate was recorded by the EEG, one side of the circuit consisting of the grid and the other side a flexible, light wire attached to a stainless steel clip located on the back of the animal near its head. The clips were attached within the first week following delivery and it was necessary to replace them about every two to three weeks.[3]

3. *Experimental Procedure*

The experiment consisted of the following successive phases:

a. Control observations. During a 21-day period each *S* was placed in the observation box for two minutes per day in each of several daily sessions and various measures were obtained (see later section).

b. Electrical stimulation. On three alternating afternoons, each *S* in the two *L* groups was given one electrical US and each *S* in the two H groups three electrical US. The *L* and *H* groups differed only in the number of stimuli per day and not in the intensity of the US. The US intensity was 150 volts, 60 cycle AC. These stimuli will be designated as the first US.

c. Extinction. The effects of the US upon the various response measures were studied in the absence of further traumatic stimulation for a 41-day period.

d. Re-stimulation. On the 41st day following the US, each *S* in the *L* groups was given one additional US and each *S* in the *H* groups three additional US. The US intensity was the same as that for the first US. These stimuli will be designated as the second US.

e. Mild stimulation. On three alternating afternoons, each *S* in the *H* groups was given a series of 10 mild electric stimuli. The intensity was set to produce a consistent but barely observable response.

f. Re-extinction. The experiment was terminated after an additional 23 days of observation.

---

[3] The procedure for recording heart rate was not entirely satisfactory. Contact potentials produced by the *S*'s movement often nullified the cardiac potentials. A much better procedure, which was discovered after the present study had begun, is that of recording between two clips, one attached near the caudal end and the other attached near the cephalic end of the spinal cord. In this procedure, movements do not interfere with the HR recording.

During the control period and each of the two extinction periods, each *S* was placed in the observation box for a period of 2 min. three times each week: Wednesday, Friday, and Sunday evenings between 8:00 p. m. and 12:00 midnight. The observation box was located in an experimental room near the animals' living quarters, and the *S*s were transported to the experimental room in a cardboard box. All the rats in a single cage were transported at one time, and the order of transportation within and between sex groups was randomized in the extinction sessions. The *S*s were placed in the observation box in the chance order that they were contacted by the experimenter as he placed his hand into the box.

4. *Measures*

The following measures were employed:

a. Motor activity. Scaled maps of the observation box were prepared. These maps divided the box into four quadrants and showed a central rectangular area consisting of the middle half of the box. The approximate path of a given *S* during an observation period was then traced with a pencil and the length of this tracing was subsequently measured by a cartometer. Three measures of motor activity were obtained by the cartometer procedure: the total length of the path, the per cent of distance in the central area of the box, and the number of quadrants entered.[4] A fourth measure of motor activity, the total duration of activity during the 2-min. interval, was obtained from the phonograph cartridge previously described. The duration of motor activity was measured in terms of the time in seconds during which a *S*'s movements resulted in a pen deflection of ⅛ in. or greater. Cartometer readings and motor recordings were obtained once each week, but the number of quadrants entered was noted in each observation period.

b. Heart rate. The heart rate (HR) in four to six 5-sec. periods during the 2-min. observation interval was averaged for each *S* and converted to minute rate. The HR was recorded once each week and could be obtained only in intervals in which a *S* was relatively quiescent. The recording technique did not permit an analysis of comparable time intervals in the different animals.

c. Food consumption. A daily record was kept of the approximate amount of food consumed per cage. Some 20 to 30 per cent of the food given was lost through the bottom of the cages, and there was no attempt to evaluate the amount lost. The food remaining in each cage was weighed at 10:00 a. m. each morning and subtracted from the amount fed. The cages were cleaned at this time, and the rats were without food until 5:00 p. m. Thus, the *S*s were satiated at the time they were placed in the observation box.

---

[4] The control period of this experiment was underway when a related study by Weininger (1956) was reported (see discussion). The cartometer, food, and weight procedures were incorporated on the basis of Weininger's findings.

d. Body weight. The *S*s were weighed five times: (*a*) on the second day of the first period of electrical stimulation; (*b*) during the middle of the first extinction period; (*c*) on the day of re-stimulation; (*d*) five days after the termination of mild stimulation; and (*e*) five days before the termination of the study.

e. Other behavioral measures. A record was kept of the number of times in each session a *S* paused and sniffed, stood up on its hind legs (standing), scratched at the box (digging), washed its face or body (preening), scratched its body, assumed stereotyped postures (posturing), shook itself (shaking), and defecated. Whether an animal did or did not urinate in the observation box was also noted.

C. RESULTS

The statistical analyses of all except the weight data were based upon three selected sessions prior to the first US, five sessions during the first extinction period, the session following the second US, and three sessions following the procedure of mild stimulation. The sessions selected for analysis were those in which all measures had been obtained. These sessions were distributed fairly evenly (about one week apart) throughout the study.

The degree of agreement between two observers on one of the three selected prestimulation sessions was determined for each of 11 measures that did not permit a completely objective method of recording or scoring. The Pearson product-moment coefficient of correlation was used for 10 of the measures and the phi-coefficient for one measure (urination). The findings summarized in Table 1 indicate statistically significant agreement between observers for each response measure, although it will be noted that urination is the least reliable measure.

1. *Motor Activity*

The effects of electrical stimulation upon two measures of motor activity are summarized in Figures 1 and 2. Median measures of motor activity are plotted as a function of time for each group of *S*s.

Figure 1 indicates that for the ML and MH groups the duration of activity during the 2-min. observation period declined precipitously after the first US. There was good evidence of recovery (extinction) in the interval between the first and second US, although there appears to have been a greater degree of recovery in the ML group than in the MH group. The second US again depressed activity, but both groups gave evidence of recovery during the ensuing re-extinction period. The general level of activity of the MH group remained below that of the ML group during the final phase of the experiment.

TABLE 1
DEGREE OF AGREEMENT BETWEEN TWO OBSERVERS FOR EACH OF 11 RESPONSE MEASURES AS DETERMINED BY THE PRODUCT-MOMENT COEFFICIENT OF CORRELATION OR BY THE PHI-COEFFICIENT

Response measure	r	rΦ	p
Defecating	1.00		$<.01$
Total length of path	.99		$<.01$
Standing	.96		$<.01$
Duration of activity	.90		$<.01$
Number of quadrants entered	.86		$<.01$
Preening	.84		$<.01$
Squeaking	.82		$<.01$
Heart rate	.82		$<.01$
Sniffing	.81		$<.01$
Digging	.72		$<.01$
Urinating		.54	$<.01$

Differences among the several groups after the introduction of electrical stimulation were tested by the binomial expansion. Each group was compared to each of the other two groups after the first US, with all nine experimental sessions considered as a unit. On the assumption that two groups being compared did not differ significantly from each other during the control period and that the stimulus presentations had no effects upon the observed behavior, then at each of the nine sessions the probability that the median value for one group would be greater than (or less than) that for the second group was one-half.

Intergroup differences during the control period were also evaluated by the binomial expansion, but the small number (three) of median values for each group precluded their use in the statistical analyses in the manner just described. Therefore, the following precedure was employed. On the first control session the five measures for the MC group and five measures for the ML group were combined into a single distribution and the median was derived. The median was also derived for the combined MC and MH groups and for the combined ML and MH groups. In each case the binomial expansion was applied on the assumption that by chance alone one-half of the measures for each group would be above and one-half below the median for the two groups under consideration. The data for the second and third control sessions were treated in a similar manner.[5]

No statistically significant differences among the three groups were found over the initial control period of the experiment. Statistical analysis of the data for the nine sessions following the first US indicated significantly shorter duration of activity for each of the two shock groups as compared to that for the control group ($p<.002$ in each case). The ML and MH groups did not differ reliably from each other, although there was a trend toward less activity in the MH group.

Similar findings were obtained for the female *S*s. As shown in Figure 1, the FL and the FH groups exhibited a marked decrease in the duration of activity after each US presentation. The females differed from the males in that the FC group also tended to be less active after the first US. No statistically significant differences were found between any two groups during the

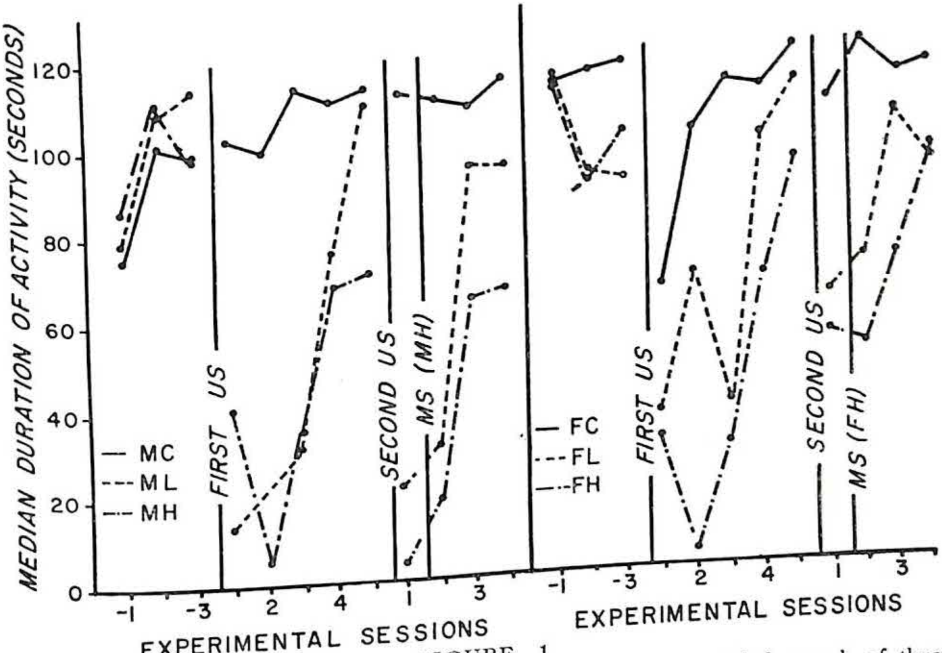

FIGURE 1

Median duration of activity for each of three male groups and for each of three female groups on each of several selected experimental sessions. In this and in Figures 2, 3, 4, and 6 the numbered sessions with minus signs attached denote control sessions; mild stimulation is abbreviated MS.

---

[5] Unless otherwise indicated, all statistical tests were based upon an expansion of the binomial. In addition, the female control and post-shock data, as well as the data for other response measures, were analyzed by procedures similar to those described.

control period, but each shock group differed reliably from the FC group and from each other during the sessions following the first US ($p < .002$ in each case). The mild shocks did not seem to have any noteworthy effects on the duration indicant for either the MH or FH group.

Figure 2 shows the effect of the two US upon the total length indicant (length of path run) for the male and for the female *S*s. During the control period, the MC group remained consistently below the ML and MH groups. While the MC group exhibited a systematic increase in the total length indicant during the course of the study, the two US groups showed a marked decrease in activity after the first US with little systematic evidence of recovery. By the end of the experiment, the two traumatized groups gave little indication of approaching either their own prestimulus levels of activity or the post-stimulus level of the control group.

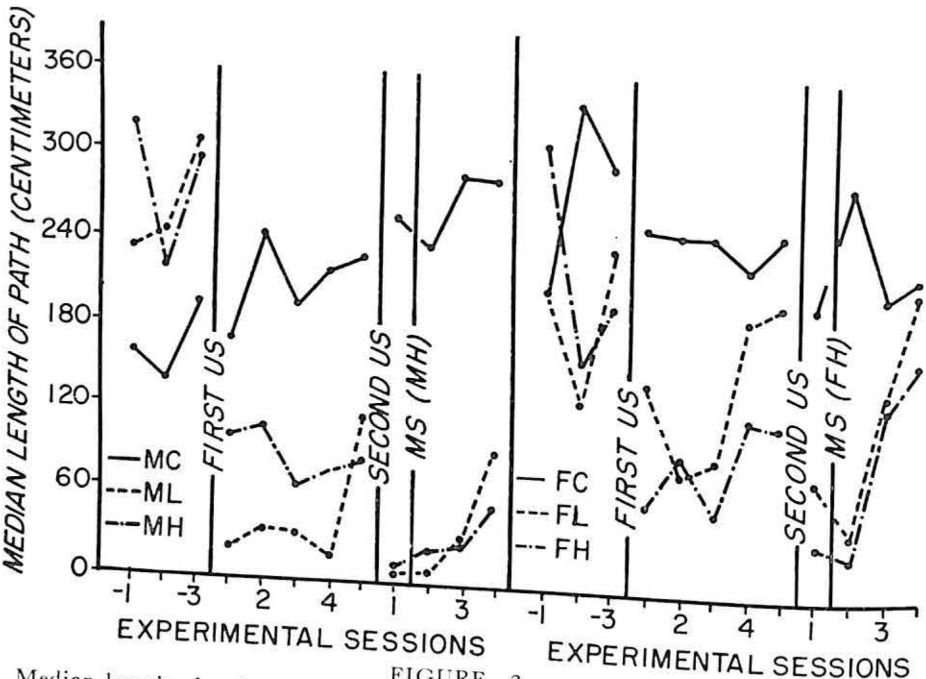

FIGURE 2

Median length of path traveled for each of three male groups and for each of three female groups on each of several selected experimental sessions. These sessions are the same as those shown in Figure 1.

No reliable differences were found among the several groups during the control period. After the introduction of electrical stimulation, the total length indicant was significantly shorter for the ML and for the MH group as compared to the control group ($p<.002$ in each case), although the two traumatized groups did not differ reliably from each other. The ML group remained below the MH group during most of the interval between the first and second US, but differences between the two curves disappear after the mild stimulation procedure. Any effect of mild stimulation upon recovery from trauma appears to be of negligible magnitude.

The findings for the female *S*s with regard to total length of path show the FC group to have been more variable than the MC group. As in the case of the males, the two US appreciably altered the activity of the FL and FH groups. The female shock groups showed a relatively greater recovery than the male shock groups. During the period following the first US the FL and FH groups were found to differ significantly from the control group ($p<.002$) in each case) and from each other ($p<.02$). No reliable intergroup differences were noted during the control period.

The decreased activity of the shocked male *S*s as compared to the male control *S*s was evident in the central distance indicant. Data for the central distance indicant have been placed on file with the ADI (see Figure A).[6]

In general, the findings for length of path in central area were similar to the total distance findings from which they were derived, although the differentiation between the shock and control groups was not clear. The pertinent aspects of these findings are as follows: (*a*) the low point of activity for each of the four shock groups was reached after the second US; (*b*) there was evidence of recovery between the first and second US in both the ML and the FL groups; (*c*) the FL and FH groups approximated the performance of the FC group by the second session following the mild stimulation procedure, i. e., 17 days after the mild stimuli; (*d*) the MH group was more active than the ML group in the last plotted session (23 days after the mild stimulation); and(e) the results for the females were less consistent than those for the males.

On the first control session the central distance indicant was significantly shorter for the MC group than for the MH group. During the period following the first US both the ML and MH groups traveled significantly shorter distances in the central area than did the MC group ($p<.002$ in each case), although there was no reliable difference between the ML and MH groups. Central distance was significantly shorter for the FH group as compared to the FL and FC groups after the first US ($p<.02$ in each case), but the FL and FC groups did not differ reliably.

---

[6] Additional data have been placed on file with the American Documentation Institute. Order Document No. 6337, remitting $1.75 for 35 mm. microfilm or $2.50 for 6 by 8 in. photocopies.

The data for number of quadrants entered also revealed that the shock *S*s differed markedly from the control *S*s (see ADI Figure B). The major findings for this measure of motor activity are as follows: (*a*) the first US produced a decrement in the number of quadrants entered by each of the shock groups; (*b*) the separation between the control group curve and each of the two shock groups curves was greater for the males than for the females; (*c*) the low point of activity for each of the four shock groups occurred after the second US; (*d*) the FL group showed some tendency to recover in the period preceding the second US; and (*e*) the FL and FH groups showed a tendency to recover during the final extinction period following the mild stimulation procedure.

No statistically significant differences were found among the male groups or among the female groups during the control period. Statistical analysis of the data for the nine sessions following the first US indicated that theML and MH groups entered significantly fewer quadrants than the MC group ($p<.002$ in each case), but that the ML group did not differ reliably from the MH group. Similarly, each of the two female shock groups entered significantly fewer quadrants than the FC group after the introduction of electrical stimulation ($p<.02$ for FL-FC comparision, $p<.002$ for FH-FC comparison), but the two female US groups did not differ significantly from each other.

As has been stated previously, the duration and total length indicants were plotted in terms of medians for each group. However, the per cent of total distance traveled in the central area was plotted in terms of means and the number of quadrants entered in terms of total frequencies for each group. With respect to these last two measures of motor activity, the position of a given point representing the behavior of a group of *S*s could have been appreciably influenced by the behavior of a single *S*. In this way, the apparent effect of electric shock upon a group of *S*s could have been a function of its consistent effect upon one rat only. However, inspection of the raw data for the two response measures revealed that, for several sessions following the first US and following the second US, three, four, or even five *S*s in each of the shock groups evidenced behavioral changes.

2. *Heart Rate*

The major trends in the plotted data for HR (see ADI Figure C) may be summarized as follows: (*a*) each of the three male groups evidenced an increase in HR during the prestimulation period and then a decrease following the first US; (*b*) each of the three female groups showed a more or less progressive decrease in HR from the beginning to the end of the experiment; (*c*) each of the two female shock groups tended to have lower HR's than the female control group during the prestimulation period; (*d*) the male shock groups tended to have higher HR's than the MC group but the female US groups tended to have lower HR's than the FC group. However, there were no reliable differences in HR among the male groups or among the female groups.

3. *Food Intake*

The main findings relevant to the effect of electric shock upon the quantity of food consumed by male *S*s and by female *S*s are presented in Figures 3 and 4, respectively. The inclusion of the data for a greater number of experimental sessions was made possible because food consumption was recorded on a daily basis, whereas the behavioral measures described previously were recorded on selected sessions only.

Examination of the several curves in Figures 3 and 4 reveals a finding of especial interest. The food consumption during presentation of the first US decreased markedly not only for the male and female shock groups but also for the male and female control groups. Further inspection of these curves

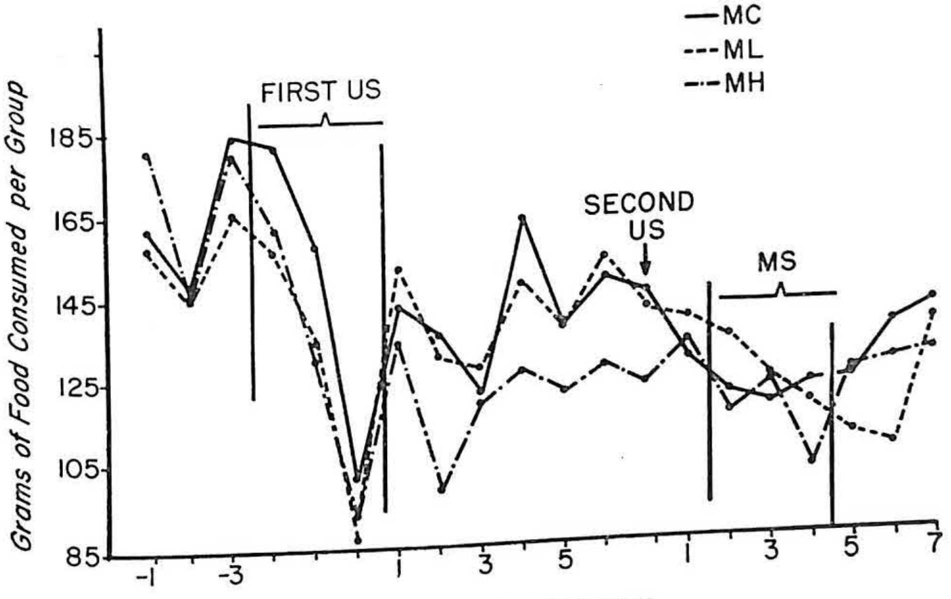

FIGURE 3

Number of Grams of Food Consumed by Each of Three Male Groups on Each of Several Selected Experimental Sessions

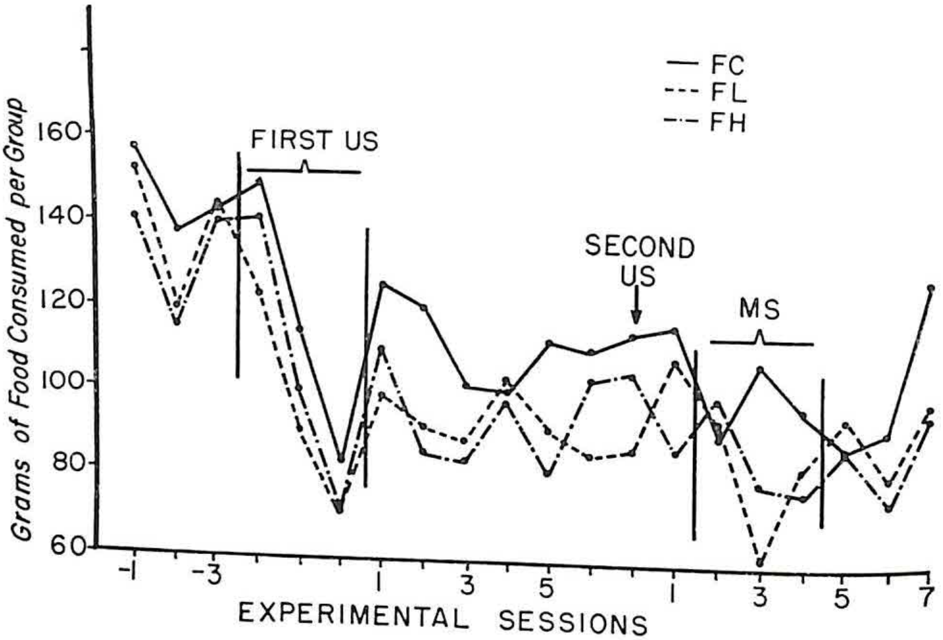

FIGURE 4

Number of grams of food consumed by each of three female groups on each of several selected experimental sessions. These sessions are the same as those shown in Figure 3.

shows little difference between the ML and MC groups in amount of food eaten during the days following the first US. The curve for the MH group remains consistently below those for the other two groups for several days following the first US, but consistent differences among the three groups disappear by the second US. The decrease in food consumption following the second US appears to have proceeded at a slower rate than that following the first US. By the end of the mild stimulation procedure the MH group evidenced a decrement in food consumption. It may also be noted that the ML and MH groups tended to recover during the terminal days of the experiment.

As compared to the curves for the males, the curves for the female shock groups tend to remain below that for the female control group following the first US. During the mild stimulation procedure both the FL and FH groups evidenced decreases in food consumption with, rather surprisingly, the FL group showing the sharper decline. The two female US groups tended to recover during the terminal days of the study, but they did not so closely approach the level of their control group as did the male shock groups.

4. *Body Weight*

In any study involving serial determinations of a rat's body weight initial weight tends to be correlated with terminal weight, regardless of the length of time between weighings. In the present investigation, for example, the Pearson product-moment coefficient of correlation between initial weight and terminal weight was 0.82 for both males and females. The effect of initial weight on terminal weight was eliminated by a regression analysis.

For the second, third, fourth, and fifth weighing periods the standard error and the best fitting linear equation (in the least squares sense) describing these subsequent weights as a function of the first weights were obtained. Average T-scores were then computed for each group of *S*s at each weighing period. The procedure was subject to the inherent limitations that linear equations derived from a *N* of 15 are relatively unstable. Nonetheless, there is reason to believe that the differences between groups aside from experimental effects would be more or less equal. Figure 5 presents the mean T-score for the various groups of *S*s for the last four weighings. Data for the first weighing do not appear because the first weighing was used in establishing the various regression relationships. Examination of the plotted values reveals

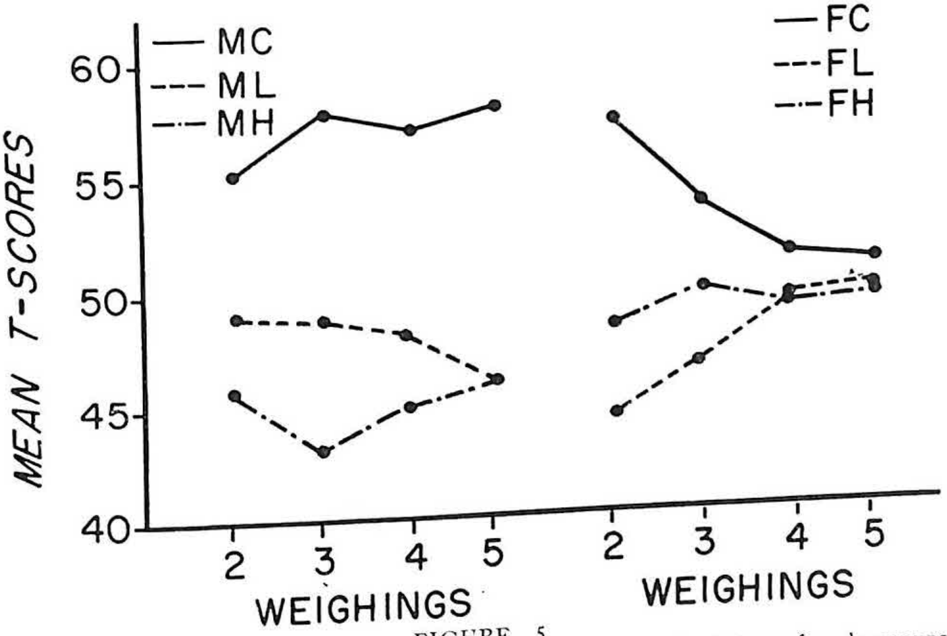

FIGURE 5

Mean T-scores for each of three male groups and for each of three female groups at each of the four weighings following the first US. Scores for the first weighing, which occurred during presentation of the first US, do not appear.

that at each of the several weighings *S*s in the MC group weighed more than would have been predicted on the basis of the regression analysis, whereas *S*s in the ML and MH groups weighed less than would have been predicted. Actual weight-predicted weight differences were greater for the MH than for the ML group, but these differences were minimized as the experiment progressed. Actual weights of *S*s in the MH group were approaching predicted weights by the end of the study, while the opposite trend was noted for *S*s in the ML group.

Actual weights of female control *S*s exceeded predicted weights over the several weighings, but magnitude of differences decreased during the terminal days of the study. Predicted weights of *S*s in each of the two shock groups exceeded actual weights. However, in contrast to the data for males, actual weight-predicted weight differences were greater for the FL than for the FH group at early weighings. Also, by the end of the study actual weights of *S*s in each of the two shock groups approached more closely the predicted weights than was the case of male shock *S*s.

5. *Other Behavioral Measures*

The graphs showing the frequency of sniffing, standing, shaking, digging, preening, scratching, posturing, and urinating have been placed on file with ADI (see Figures D through K). Figure 6 presents the data for defecation

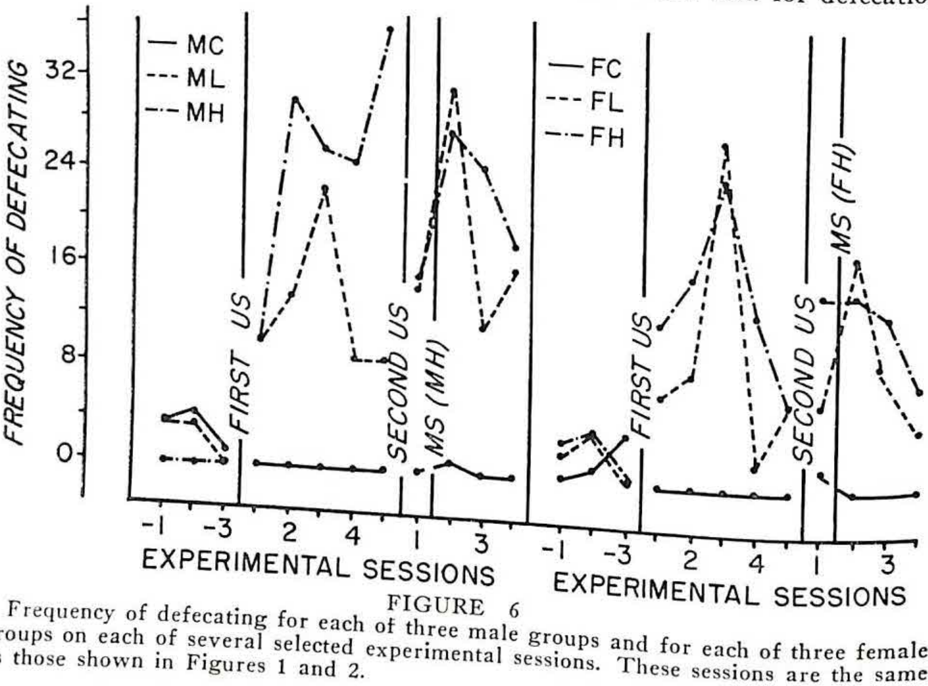

FIGURE 6

Frequency of defecating for each of three male groups and for each of three female groups on each of several selected experimental sessions. These sessions are the same as those shown in Figures 1 and 2.

during the several experimental sessions. It may readily be seen that following first US the female shock groups extinguished more rapidly than the male shock groups, the frequency of defecation for both males and females increased after the second US, and the mild stimulation procedure appeared to produce no obvious effect upon frequency of defecation in the MH and FH groups. While the data will not be presented in this paper, the US groups, as compared to the control groups, had a greater frequency of defecation on the way to and from the experimental room.

For each of the nine response measures previously mentioned the results of the statistical analysis of the data for the nine sessions following the first US are presented in Table 2. Examination of the table reveals two findings of particular significance: (*a*) for most of the several response measures the *L* and *H* groups differed reliably from the *C* groups, although the *L* and *H* groups failed to differ from each other; and (*b*) the differences

TABLE 2

Binomial Test of Differences in Several Response Measures between Male Experimental Groups and between Female Experimental Groups During Period Following First US

Response measure	Males	p	Females	p
Sniffing	MC vs. ML	$< .02$	FC vs. FL	$< .002$
	MC vs. MH	$< .02$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Standing	MC vs. ML	$< .002$	FC vs. FL	$> .05$
	MC vs. MH	$< .002$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$< .002$
Shaking	MC vs. ML	$< .002$	FC vs. FL	$> .05$
	MC vs. MH	$< .002$	FC vs. FH	$> .05$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Preening	MC vs. ML	$< .02$	FC vs. FL	$< .002$
	MC vs. MH	$> .05$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Scratching	MC vs. ML	$< .002$	FC vs. FL	$> .05$
	MC vs. MH	$< .002$	FC vs. FH	$> .05$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Posturing	MC vs. ML	$< .02$	FC vs. FL	$> .05$
	MC vs. MH	$< .02$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Digging	MC vs. ML	$> .05$	FC vs. FL	$> .05$
	MC vs. MH	$> .05$	FC vs. FH	$> .05$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Urinating	MC vs. ML	$> .05$	FC vs. FL	$> .05$
	MC vs. MH	$> .05$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$
Defecating	MC vs. ML	$< .002$	FC vs. FL	$< .002$
	MC vs. MH	$< .002$	FC vs. FH	$< .002$
	ML vs. MH	$> .05$	FL vs. FH	$> .05$

between shock and control groups were more pronounced for the males than for the females. None of the behavioral measures revealed reliable intergroup differences during the control period.

Since for each of these measures total frequencies per group were plotted as a function of experimental sessions, the question again arises as to whether electrical stimulation had an effect upon groups of *S*s or upon a single *S* only within a shock group. Inspection of the raw data revealed group effects following each of the two shock series. For some measures, however, the recovery process appeared to proceed at differential rates in animals of a shock group, so that by the later phases of the extinction periods perhaps only one or two *S*s were reflecting the effect of electric shock.

6. *Additional Analyses*

In order to compare the several behavioral measures in terms of duration of disturbance (or recovery from trauma) during the interval between the first and second US, the data in Figures 1, 2, 6, and in ADI Figures *A*, *B*, and *D* through *K* were considered in the following manner. For a given behavioral measure and for a given shock group two points were noted: Point *A*—that corresponding to the greatest change in response magnitude (as compared to prestimulation performance) during the five sessions representative of the 41-day period between the first and second US; Point *B*—that representing the performance of the appropriate control group on the fifth and last selected session (Session 5) of this period between US presentations. The vertical distance between these two points was then measured and divided into four equal segments. By treating Point *B* as the theoretical point of maximal recovery for the shock group in question, the actual point of maximal recovery following Point *A* was assigned a numerical value of 1, 2, 3, or 4, depending upon the particular segment in which it fell. A rating of 2 indicated a greater degree of recovery than a rating of 1, and so on.

In addition to rating on a 4-point scale the degree of maximal recovery, determination was also made of the number of days succeeding the first US that were required to achieve 50 per cent recovery. For a given response measure and for a given shock group the first point following Point *A* that could be assigned a rating of at least 3 was noted. Then a record was made of the number of days between the termination of the first US and the experimental session on which this noted point was reached.

The procedures described in the previous sections may be illustrated through the use of a specific example—the performance of the ML group with respect to one measure of motor activity, the duration of activity. Examination of Figure 1 reveals that Point *A* was reached on Session 1 following the first US. Progressive recovery from trauma occurred over the succeeding sessions until the performance of the ML group most closely approached that of the MC group on Session 5. The point representing the performance of the ML group on Session 5 was accorded a rating of 4 on the 4-point scale of recovery. The first point for the ML group to receive a rating of 3 (indicating at least 50 per cent recovery) was reached by Session 4; the number of days between the end of the first US and Session 4 was 28. The values 4 and 28 together with comparable values for each of the three remaining shock groups with respect to each of 12 other response measures are summarized in Table 3.

TABLE 3

DATA ON RECOVERY FROM EFFECTS OF ELECTRICAL STIMULATION FOR EACH OF FOUR SHOCK GROUPS WITH RESPECT TO EACH OF SEVERAL RESPONSE MEASURES

Response measures	Shock groups ML	MH	FL	FH
Duration of activity	4/ 28 4/ 17	3/ 28 3/ 17	4/ 28 4/ 17	4/ 28 3/ 23
Total length of path	2/> 41 2/> 41	1/> 41 1/> 41	3/ 28 4/ 17	2/> 41 3/ 17
Length of path in central area	3/ 35 2/> 41	1/> 41 3/ 23	2/> 41 4/ 17	1/> 41 3/ 17
Number of quadrants entered	2/> 41 1/> 41	1/> 41 1/> 41	3/ 28 2/> 41	1/> 41 3/ 17
Sniffing	2/> 41 2/> 41	2/> 41 1/> 41	2/> 41 2/> 41	2/> 41 2/> 41
Standing	4/ 35 3/ 23	2/> 41 1/> 41	4/ 19 4/ 23	3/ 35 3/ 23
Shaking	1/> 41 1/> 41	3/ 35 1/> 41	4/ 19 3/ 23	1/> 41 2/> 41
Preening	3/ 35 1/> 41	4/ 14 1/> 41	3/ 28 3/ 17	3/ 35 1/> 41
Scratching	1/> 41 1/> 41	3/ 35 1/> 41	2/> 41 1/> 41	2/> 41 1/> 41
Posturing	4/ 19 4/ 2	4/ 19 4/ 17	4/ 19 4/ 2	4/ 19 4/ 2
Digging	4/ 28 4/ 23	3/ 35 4/ 2	1/> 41 3/ 23	1/> 41 4/ 2
Urinating	4/ 19 4/ 2	4/ 19 4/ 2	4/ 14 4/ 2	3/ 35 4/ 17
Defecating	3/ 28 3/ 17	1/> 41 2/> 41	4/ 28 3/ 23	3/ 35 3/ 23

Note.—For each response measure and for each shock group the two numbers in the first line represent, respectively, the degree of recovery on a 4-point scale and the number of days to 50 per cent recovery for the 41-day period following the first US; the two numbers in the second line represent the comparable values for the 23-day period following the mild stimulation procedure.

In order to ascertain the effect, if any, of mild stimulation upon duration of disturbance and the recovery process, procedures similar to those previously described were used. For a given response measure and for a given shock group Point *A* corresponded to the greatest change in response magnitude (as compared to the point of maximal recovery occurring in the first US—second US· interval) during the four sessions representative of the 31-day period following the second US. Point *B* represented the performance of the appropriate control group on the last selected session, Session 4, of this period. The actual point of maximal recovery for a particular shock group was given a rating of 1, 2, 3, or 4 in the manner described previously. The number of days to 50 per cent recovery was based upon the 23-day period following the mild stimulation procedure. The several values so derived for each of the four shock groups and for each of the 13 response measures are also presented in Table 3.

Inspection of Table 3 reveals that some of the behavioral measures were characterized by a lesser degree of recovery, according to the criteria used, than were others. All four shock groups showed little recovery with respect to sniffing and scratching. In addition, the male shock groups exhibited a fairly prolonged duration of disturbance when total length of path, number of quadrants entered, and shaking are considered.

Although the mild stimulation procedure was apparently ineffectual in alleviating pathological disturbances in the *H* groups, it should be remembered that data were plotted for selected experimental sessions only; i.e., those sessions on which data for all measures were available. The three selected sessions following the series of mild stimuli may have failed to provide for a sufficiently discriminative evaluation of the effect of these stimuli upon recovery from trauma. However, data for additional sessions were available for each of five response measures—number of quadrants entered, squeaking (both when animals were put into and taken out of the observation box), urinating, defecating, and food consumption. With respect to each of these five measures statistical comparisons between the male and between the female shock groups were made on the basis of performances (*a*) on seven sessions prior to the second US and (*b*) on 10 sessions following the mild stimulation procedure. For both (*a*) and (*b*) the intersession intervals were two to three days in duration.

Application of the median test to the data for the several groups yielded the results summarized in Table 4. Although five of the 24 Chi-square values indicated reliable differences between *L* and *H* groups, there was no evidence that mild shock hastened recovery from trauma. Over the seven days prior to the second US *S*s in the MH group entered a significantly greater number of quadrants, defecated a significantly greater number of times, and consumed significantly less food than *S*s in the ML group. During both experimental phases (prior to the second US and following the mild stimulation procedure) the number of *S*s that urinated was reliably greater for the FH group than for the FL group.

TABLE 4

Median Test Analysis of Differences between L and H Groups Before and After Mild Stimulation with Respect to Each of Five Response Measures

Response measure	Groups compared	Chi-square Value	
		Before	After
Number of quadrants entered	ML vs. MH	4.57*	.88
	FL vs. FH	.29	0
Squeaking (In)	ML vs. MH	.28	.21
	FL vs. FH	0	.24
Squeaking (Out)	ML vs. MH	.58	.56
	FL vs. FH	.28	.56
Urinating	ML vs. MH	0	1.88
	FL vs. FH	4.57*	10.21**
Defecating	ML vs. MH	4.57*	.20
	FL vs. FH	1.14	0
Food intake	ML vs. MH	10.28**	1.80
	FL vs. FH	0	.20

* Significant at the .05 level of confidence.
** Significant at the .01 level of confidence.

D. Discussion

The present investigation had two major purposes: (*a*) to study the effect of traumatic electrical stimulation upon the growth and behaviors of the rat; and (*b*) to study the role of mild electrical stimulation in alleviating behavioral disturbances produced by traumatic stimulation. With respect to the first of these two purposes, it was found that electric shock given over a period of three days had a reliable effect upon a number of response measures. Heart rate only failed to reflect any sustained changes attributable to stimulation. Unlike the dog reported in the previous study by Dykman and Gantt (1960), the rats exhibited no evidence of autokinesis; observed symptoms of disturbance were present immediately after shock and abated as a function of time. With respect to the second purpose, the results were negative. It was found that mild stimulation did not hasten the recovery process.

The conventional method of stimulating rats by means of a grid is not entirely adequate. The voltage transmitted depends upon the type of contact the *S* has with the grid at the time of stimulation. Variations in contact potentials would have a greater effect upon low levels of stimulus intensity than upon high levels. With weak stimuli it is difficult to adjust the voltage to produce consistently a barely observable response. Thus, in the present study, the mild stimuli may not have been sufficiently mild. Stimuli delivered through permanently implanted electrodes would be more easily regulated and would perhaps lead to more consistent results at both high and low levels of intensity.

Although, for most of the functions studied, male and female shock groups differed reliably from their respective control groups following the first US, differences between male shock groups and between female shock groups were not significant. It will be recalled that a greater number of shocks was given to *H* groups than to *L* groups but that stimulus intensity was held constant (except for variations in contact potentials). Reliable response differences between the *L* and *H* groups might have been obtained by a procedure in which the difference in the number of shocks given the groups was greater than in the present study. The time interval between shocks comprising the first US (two days in the present investigation) might also be a variable relevant to observed differences between *L* and *H* groups. However, a study by Peters and Finch (in press) showed that the duration and magnitude of behavioral disturbances in the rat covaried with the intensity of stimulation.

The several functions studied varied with respect to the duration of disturbance following trauma. Sniffing and scratching showed least evidence of recovery for both males and females during the extinction and re-extinction periods. There were a number of measures that failed to recover to the level of the control group before the administration of the second US; total length of path traveled, length of path traveled in the central area of the box, number of quadrants entered. Measures that showed appreciable recovery prior to the second US included duration of activity, standing, preening, posturing, urinating and defecating.

The precise determination of the end-points of the recovery process for all the various behaviors considered would constitute an interesting problem for future research. Once norms are established for each measure, investigations could focus upon conditions that intensify or attenuate the recovery process—drugs, social treatments, and neurological procedures.

Studies with rats by Hunt and Brady (1951), Brady (1951), Brady and Hunt (1951), and Brady (1952) showed that electroconvulsive shock diminished or eliminated a previously established conditional emotional response of crouching, defecating, or cessation of lever-pressing in a Skinner box situation. The point to be emphasized here is that the use of electroconvulsive shock was one method of alleviating experimentally-produced behavioral disturbances.

The study by Peters and Finch mentioned previously was concerned with the effect of a "social" rat on the activity of rats that already had been exposed to traumatic stimulation. The "social" rat facilitated the activity of the stressed more than the control animals.

While female *S*s evidenced patterns of response to traumatic stimulation similar to those for male *S*s, the two US presentations tended to have somewhat more pronounced effects upon male activity than upon female activity. For example, the females appeared to recover more quickly with respect to total length of path traveled, number of quadrants entered, standing, shaking, and defecating. A factor to be considered in comparing data for males and females is the female estrus cycle and its demonstrated effect on general behavior (Slonaker, 1924; Wang, 1923; and Wang, 1924). If an experimenter desires for certain purposes to produce the most clear-cut behavioral disturbances possible, then he should perhaps confine his efforts to male rats only. In any event, further study of sex differences in relation to resistance to stress would be of interest.

The decrease in the amount of food eaten per day by the shock groups following traumatic stimulation was expected. It was surprising, however, to find that food consumption during presentation of the first US was also decreased in the male and female control groups. This finding may have been the result of some mistake made in caring for the animals or possibly an unseasonably hot day. The animal room was not air conditioned and rats deprived of water will tend to stop eating. However, it is doubtful that the effects observed in the control groups in this study were completely artificial, since Peters and Finch in the study referred to previously observed a similar but less pronounced reduction in gnawing activity and in weight of his control *Ss*. His animals were housed in air conditioned quarters.

Under the conditions of the present study, there were at least three ways in which control *S*s might have been stimulated by the sight, smell, and sounds of the stimulated *S*s: all *S*s were carried to the experimental room in the same carrying box; all *S*s were placed in the same observation box; and while the different experimental groups were housed in different cages, all cages were located on one shelf of a small animal rack. It will be recalled that the order in which *S*s were transported to the experimental room and the order in which they were placed in the observation box was randomized in the different extinction sessions. These results suggest some interesting experiments concerned with the possible communication of emotional states in rats.

An important consideration in the evaluation of the results from the study is the age of the rats at the time of traumatic stimulation. Some investigations (Scott, Fredericson, and Fuller, 1951; Scott and Marston, 1950) have shown that there are critical periods in the life of an animal during which a given experience will have a greater impact on behavior than at earlier or later times. It is also likely that these critical periods vary with different behaviors (Lorenz, 1957, pp. 83-128, pp. 288-310). Peters and Finch, in the study mentioned previously, found that shocks of relatively high intensity given to male rats at the age of 14 months permanently depressed motor activity. The determination of the critical periods for at least some of the functions investigated in the present study would constitute a significant research endeavor.

A study by Weininger (1956) published during the preliminary control phase of the present investigation presented findings relevant to those from the present study. Weininger was concerned with the effects of gentling introduced immediately after weaning upon the behavior and growth of rats. When both gentled and nongentled *S*s were observed in an open field situation, the gentled *S*s were appreciably more active than the nongentled *S*s The gentled rats showed a significantly greater mean body weight than the nongentled animals, but the latter group exhibited a greater mean weight of feces. There was no reliable difference between the gentled and nongentled animals in food consumption. After exposure to an acute stress experience, the rats were sacrificed so that their internal organs might be examined. The gentled animals exhibited less heart damage and fewer bleeding points in the stomach and duodenum than the nongentled animals, and the adrenal glands of the nongentled group were heavier than those of the other group. The present study differs from Weininger's study in terms of type of stimulation and the age of animals at the time of stimulation. The two investigations complement each other by showing that those measures common to both studies are affected in an opposite manner by gentling and traumatic stimulation.

The basic aim of the type of research presented in this paper lies in the production of behavioral disturbances in the rat under controlled conditions The technique reported here is promising in that disturbances can be produced simply and studied under a variety of experimental conditions. While it is dangerous to generalize from the behavior of lower organisms to that of humans, it is possible that refinements of the method used in this study will provide insight concerning some of the basic mechanisms underlying psychopathology and therapy.

E. SUMMARY

The purposes of the present study were: (*a*) to study the effect of traumatic electrical stimuli upon the multiple behaviors and the development of the albino rat; (*b*) to compare the various behaviors in terms of duration of disturbance; and (*c*) to study the hypothesized "therapeutic" role of mild electrical stimuli in alleviating behavioral disturbances.

Thirty albino rats, 15 male and 15 female, served as subjects. At the beginning of the study the males were two months old and the females four months old. The 15 *S*s of each sex were assigned at random to three experimental groups with five rats in each group. One group of males and one group of females comprised two one-shock per day groups; a second group of males and a second group of females comprised two three-shock per day groups; the third group of males and the third group of females served as controls. The intensity of electric shock (150 volts, 60 cycles AC) was held constant for all animals and the several groups differed only in the number of shocks per day.

The entire experiment was divided into six periods as follows: (*a*) a control period; (*b*) a period of electrical stimulation—on each of three alternate days each *S* in each of the four shock groups was given the appropriate number of shocks (one or three); (*c*) an extinction period, during which time the effects of stimulation were noted; (*d*) a re-stimulation period—on one day only each *S* in each shock group was given the appropriate number of shocks; (*e*) a "re-assurance" period— on each of three alternate days each *S* in each of the two three-shock per day groups was given a series of 10 mild electric shocks; and (*f*) a re-extinction period, during which time the behaviors of all *S*s were observed without any further administration of electrical stimulation.

Tests were conducted in an open-field situation consisting of an observation box divided into quadrants. During a given 2-min. observation period recordings were made of the following behavioral measures: duration of activity, total length of path traveled, length of path in central area, number of quadrants entered; the frequency of sniffing, standing, digging at the box, preening, stereotyped posturing, scratching, shaking, urinating, and defecating. Note was also made of heart rate, daily food intake, and changes in body weight.

The main findings from the present investigation may be summarized as follows.

1. With the exception of one measure (digging), administration of

electric shock had a reliable overall effect upon the remaining behavioral measures employed. There was a tendency toward increased digging following stimulation.

2. While within a sex grouping overall differences between shock and control groups were statistically significant, differences between one-shock per day and three-shock per day groups were not reliable.

3. The several behavioral measures differed in terms of duration of disturbance (recovery) following the first period of stimulation and the series of mild shocks. During both the extinction and re-extinction periods sniffing and scratching evidenced least recovery for both male and female shock groups. Measures showing little recovery prior to the re-stimulation period included total length of path traveled, length of path in central area, and number of quadrants entered. The other measures, such as duration of activity, standing, preening, posturing, urinating, and defecating, evidenced appreciable recovery prior to re-stimulation.

4. Electric shock appeared to have somewhat more pronounced effects upon male activity than upon female activity. During the extinction and re-extinction periods female *S*s tended to recovery more quickly than males with respect to total length of path, number of quadrants entered, standing, shaking, and defecating.

5. Administration of mild electric shocks to the two three-shock per day groups did not hasten recovery from trauma.

6. Average heart rate for each group tended to decrease progressively over the course of the study. No reliable differences in heart rate among the several groups were noted.

7. During the first period of stimulation food consumption decreased not only for shock groups but also for control groups. Throughout the course of the study female shock groups were more clearly differentiated from the female control group than were male shock groups from the male control group. There was a tendency toward a greater degree of recovery for male shock groups than for female shock groups.

8. At each of several successive weighings during the course of the study actual weights of *S*s in each of the two male shock groups were less than predicted weights, although differences between actual and predicted weight tended to be greater for the three-shock per day group than for the one- shock per day proup. At early weighings *S*s in the female shock groups also tended to weigh less than would have been predicted, with actual weight-predicted weight differences being greater for the once-shock per day group than for the three-shock per day group. As the study progressed, however, the

actual weights of *Ss* in the female shock groups approached more closely the predicted weights than did those of *Ss* in the male shock groups.

The present investigation was regarded as an initial step in the development of a standardized test procedure to be employed in studying the effects of various types of "therapies" and intracranial stimulation upon experimentally-produced behavioral disturbances. Implications for future research were discussed.

REFERENCES

1. BRADY, J. V. The effect of electro-convulsive shock on a conditioned emotional response: The permanence of the effect. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 507-511.
2. ———. The effect of electro-convulsive shock on a conditioned emotional response: The significance of the interval between the emotional conditioning and the electro-convulsive shock. *J. Comp. & Physiol. Psychol.*, 1952, **45**, 9-13.
3. BRADY, J. V., & HUNT, H. F. A further demonstration of the effects of electro-convulsive shock on a conditioned emotional response. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 204-209.
4. DYKMAN, R. A., & GANTT, W. H. A case of experimental neurosis and recovery in relation to the orienting response. *J. of Psychol.*, 1960, **50**, 105-110.
5. GANTT, W. H. Principles of nervous breakdown-schizokinesis and autokinesis. *Ann. N. Y. Acad. Sci.*, 1953, **56**, 143-163.
6. HUNT, H. F., & BRADY, J. V. Some effects of electro-convulsive shock on a conditioned emotional response ("anxiety"). *J. Comp. & Physiol. Psychol.*, 1951, **44**, 88-98.
7. LORENZ, K. Companionship in birdlife. *In* C. H. Schiller (*Ed. & Trans.*), *Instinctive Behavior.* New York: Int. Univ. Press, 1957. [Pp. 83-128. (a)]
8. ———. The past twelve years in the comparative study of behavior. *In* C. H. Schiller (*Ed. & Trans.*), Instinctive Behavior. New York: Int. Univ. Press, 1957. [Pp. 288-310. (b)]
9. SCOTT, J. P., FREDERICSON, E., & FULLER, J. L. Experimental exploration of the critical period hypothesis. *Personality*, 1951, **1**, 162-183.
10. SCOTT, J. P., & MARSTON, M. V. Critical periods affecting the development of normal and mal-adjustive social behavior of puppies. *J. Genet. Psychol.*, 1950, **77**, 25-60.
11. SLONAKER, J. R. The effects of pubescence, oestration, and menopause on the voluntary activity of the albino rat. *Amer. J. Physiol.*, 1924, **68**, 294-315.
12. WANG, G. H. Relation between "spontaneous" activity and oestrus cycle in the white rat. *Comp. Psychol. Monog.*, 1923, **2**, No. 6.
13. ———. A sexual activity rhythm in the female rat. *Amer. Nat.*, 1924, **58**, 36-42.
14. WEININGER, O. The effects of early experience on behavior and growth characteristics. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 1-9.

Laboratory for Behavioral Research
Medical Center
University of Arkansas
Little Rock, Arkansas

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 253-256.

A CRITICAL NOTE ON BARBER'S CASE-STUDY ON "SUBJECT J"*

Studies in Hypnosis Project, Massachusetts Mental Health Center

RONALD E. SHOR AND JOEL SCHATZ[1]

In a recent paper Barber (1957) draws an important generalization from the case-study of one individual referred to as "Subject *J*."

Barber describes *J* as an individual who was once a classic hetero-hypnotic somnambulistic subject who at present is a poor hetero-hypnotic subject because "he is not willing to accept an operator's statements as valid descriptions of what is or is not occurring." Barber's description however implies not only an unwillingness, but, indeed, an inability on *J*'s part to comply with the hypnotist's directions.

> Subject *J* is a 20-year-old college student who can enter trance readily and can easily induce autohypnotic phenomena in himself. However, he is not, at present, a good (hetero) hypnotic subject; he is not willing to accept an operator's statements as valid descriptions of what is or is not occurring. However, when *J* was first hypnotized and for a number of months thereafter, he behaved in a classic somnambulistic manner.
>
> Although at present (six months later) *J* can still enter trance quickly and easily and can still produce autohypnotic phenomena, *he is no longer a good subject in an interpersonal hypnotic relationship.* Although he was formerly ready and willing to accept an operator's statement that he was six years old, he now states, in the same situation, 'I know I'm not six years old, but I'll pretend that I am if you really want me to do so' . . . . . (p. 301; parentheses his; italics ours).

Barber then quotes *J* to the effect that in the beginning of *J*'s hypnotic experiences *J* did not realize he could disobey the hypnotist and consequently literally believed everything the hypnotist said.

> 'When I first had trance experience I didn't know I could disobey. I thought that anything the hypnotist said must be so. When you told me I wouldn't remember anything I really thought that I couldn't remember. When you told me that I was becoming thirsty, I didn't even stop to think that I didn't have to get thirsty. I really believed that I was going to die of thirst' (p. 301).

---

* Received in the Editorial Office on June 30, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by postdoctoral fellowship MF-7860-C2 from the National Institute of Mental Health, Public Health Service.

Barber then attempts to explain why *J* no longer literally believes the hypnotist's statements.

> Having read widely in the literature of hypnosis and having completed a university course on the subject, he had concluded that the hypnotist does not possess any special 'power' or 'ability', that a subject induces the phenomena himself, and that a subject does not have to accept the hypnotist's statements as true and valid. *As a result he can no longer experience age-regression, since he is always aware that he is not really 'six years old'; he no longer has amnesia for the hypnosis because he knows that he can remember; he does not carry out post-hypnotic suggestions because he knows that he does not have to do so* (p. 302; italics ours).

It is clear that Barber believes *J* is now unable to respond hetero-hypnotically because *J*'s increased knowledge about hypnosis stands in the way. If Barber's conclusion is correct, it has wide implications for understanding the nature of hypnosis. At least three extremely far-reaching theoretical deductions appear to be implied in Barber's case-study:

I. Only those people who believe that they must obey a hypnotist in order to remain in trance can be good hetero-hypnotic subjects.

II. Once a person realizes that trance need not imply obeying someone else, the possibility is thereby closed for him to follow suggestions from an operator except as sheer pretense and without perceptual alterations.

III. Hetero-suggestibility in trance is based on ignorance, and knowledge necessarily and irreversibly prevents its occurrence.

The present writers feel they must take firm exception to Barber's observations. The implications of Barber's report are so far-reaching and of such potential importance if correct that the writers feel they cannot allow the report to stand as is in the literature without critical comment. Although the present writers share the view that *J* was unwilling to respond hetero-hypnotically during these few occasions with Barber (and yet entered an autohypnotic state) they are firmly opposed to the interpretation that this means *J* was unable to respond hetero-hypnotically, either at that time or in the future.

In order to put Barber's interpretation to a direct test the senior author induced hypnosis in Subject *J*, the co-author of this report. In setting up this experiment with *J*, the senior author discussed the previously published report as well as *J*'s feelings about entering hetero-hypnosis in Barber's study. Every effort was made to elicit *J*'s interest and cooperation. As *J* made a genuine intellectual contribution in helping to comprehend these issues and since his subjective reports were vital he agreed to take responsibility for them in the literature, by name, as a co-author.

During hypnosis loss of awareness of self and environment was first induced, then an age-regression, and finally Subject *J* was awakened with the suggestion of total amnesia until a pre-arranged cue word was spoken by the operator.

Upon awakening it was observed that the amnesia was complete. Subject *J* was told about the amnesia, the age-regression, and about the cue word, but this knowledge did not result in the breakdown of the amnesia.

The following day the operator spoke the cue word to Subject *J* and immediate recall followed. In describing the experience of age-regression *J* reports that phenomenologically it was totally real to him. *J* further reports that these occurrences were as profound a hetero-hypnotic experience as he has ever had, and that once the decision to permit himself to enter into hetero-hypnosis with the operator was made he felt unable to rescind the decision; he felt compelled to experience reality in terms of the operator's instructions.

It is evident that Subject *J* is still able to function as "a classic hetero-hypnotic somnambulistic subject" under some circumstances. Thus the implication must be rejected that increased knowledge about hypnosis makes impossible the induction of a typical hetero-hypnotic state.

Some light may be thrown on factors underlying Barber's descriptions by the following considerations not discussed in Barber's report. (*a*) The senior author was the instructor of the university tutorial on hypnotis attended by Subject *J*. (*b*) In this seminar *J* *specifically* had been taught to be able to produce profound auto-suggestive phenomena in himself while retaining a choice of whether or not to respond to hetero-suggestions. (*c*) In the sessions with Barber *J* reports that an insufficient effort was made to elicit his wish to enter a hetero-hypnotic relationship with the experimenter. (*d*) Subject *J* reports having felt mild hostility toward Barber at the time.

Depending upon one's theoretical predilections one may choose to interpret *J*'s failure to respond with Barber in a variety of ways. One may emphasize *J*'s unwillingness to refrain from critical self-examination; his lack of commitment to be responsive to Barber's suggestions; or one might argue that in training *J* to experience auto-hypnosis the senior author had implicitly implanted 'locking suggestions' which applied when Barber worked with *J*. But how ever the reader may choose to interpret the reason for these occurrences it is clear that Subject *J* retained his ability to enter "classical hypnosis" when provided with the appropriate situation—despite his increased knowledge about the phenomenon.

Summary. A critique is given of Barber's case-study of a previously excellent somnambulist who allegedly is no longer able to respond in a hetero-hypnotic situation. An experiment was performed by the senior author to show that Barber's subject, the co-author, is still able to respond in the classic hetero-hypnotic somnambulistic manner.

REFERENCE

1. BARBER, T. X. Hypnosis as perceptual-cognitive restructuring: III. From somnambulism to autohypnosis. *J. of Psychol.*, 1957, **44**, 299-304.

Massachusetts Mental Health Center
Studies in Hypnosis Project
74 Fenwood Road
Boston 15, Massachusetts

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 257-269.

THE PERCEPTION OF A DEPTH INTERVAL WITH BINOCULAR DISPARITY CUES*

US Army Medical Research Laboratory, Fort Knox, Kentucky

WALTER C. GOGEL[1]

A. PROBLEM

This study will investigate the perception of a depth interval at different distances from *O,* when this depth interval is supported only by binocular disparity cues. More specifically, the problem of this study is to determine how this physical depth interval must vary in order to remain perceptually unchanged at different distances from *O.*

The problem can be illustrated with the aid of Figure 1. In the upper portion of Figure 1, two depth intervals X_{ef} and X_{gh} are represented. The eyes of *O* are shown at the left with the distance between the pupils of the eyes labeled *A*. In the lower portion of Figure 1, the angular subtense of the interpupillary distance *A* at the distances D_e, D_f, D_g, and D_h is shown as α_e, α_f, α_g, and α_h. The binocular disparity associated with X_{ef} is $\alpha_e - \alpha_f$ and the binocular disparity associated with X_{gh} is $\alpha_g - \alpha_h$. It can be readily demonstrated that the relation between binocular disparity and the physical depth interval *X* is

$$\alpha_e - \alpha_f = \frac{A\, X_{ef}}{D_e^2 + D_e\, X_{ef}} \qquad (1),$$

with a similar equation applying to $\alpha_g - \alpha_h$ etc.

It is clear from Equation 1, that if two depth intervals at different distances (*D*) have the same linear extent (*X*), they cannot produce equal binocular disparities. Thus, the binocular disparity stimulus resulting from a constant depth interval is not the same at different distances. The problem is to determine how this affects the perceptual response or, more generally, how the perceived size of a depth interval varies as a function of both the distance (*D*) of the interval from *O* and the magnitude of the binocular disparity. To investigate this problem, *O* will have the task of adjusting depth intervals such as X_{gh} to be perceptually equal to X_{ef}. The resulting change in the required magnitude of X_{gh} will be examined as a function of *D*.

---

* Received in the Editorial Office on July 5, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank Richard E. Granda for his help in collecting and analyzing the data of this study.

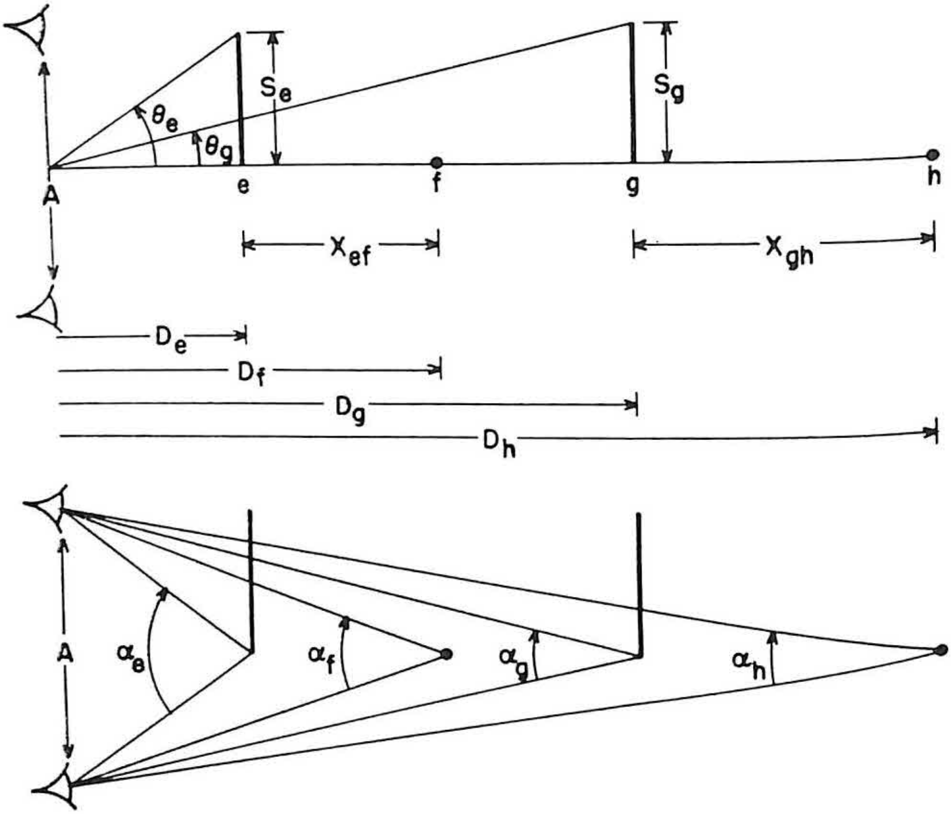

FIGURE 1
A Schematic Diagram for Considering the Binocular Judgment of Depth Intervals

It was required in this experiment that no relative depth cues except binocular disparity be present to determine the perceived depth intervals. This was the only restriction with respect to available cues which was placed on the study. As will be discussed later, any number of cues to the absolute distance of the depth interval (its distance from *O*) can be present. It will be demonstrated that under these conditions, the resulting responses can be predicted from a theoretical point of view which emphasizes the perceptual interrelation of frontal and depth extents. This point of view requires that the change in the perceived frontal size of an object as a function of *D* must be determined before the similar function with respect to the perceived depth interval can be predicted (5). For this reason, frontal judgements as well as depth judgements were made by *O*. The notation for considering the results from the frontal judgements is illustrated in Figure 1 in which the two frontal extents S_e and S_g are shown to subtend the angular extents θ_e and θ_g respectively. The task of *O* was to judge when S_e was the same size as S_g. It is asserted that the results of this frontal judgement are important for predicting the magnitude of X_{gh} which will appear equal to X_{ef}.

B. METHOD

1. *Apparatus*

a. The Visual Alley. A schematic perspective diagram of the visual alley used in the experiment is shown in Figure 2. In this figure, Objects *e, g, i,*

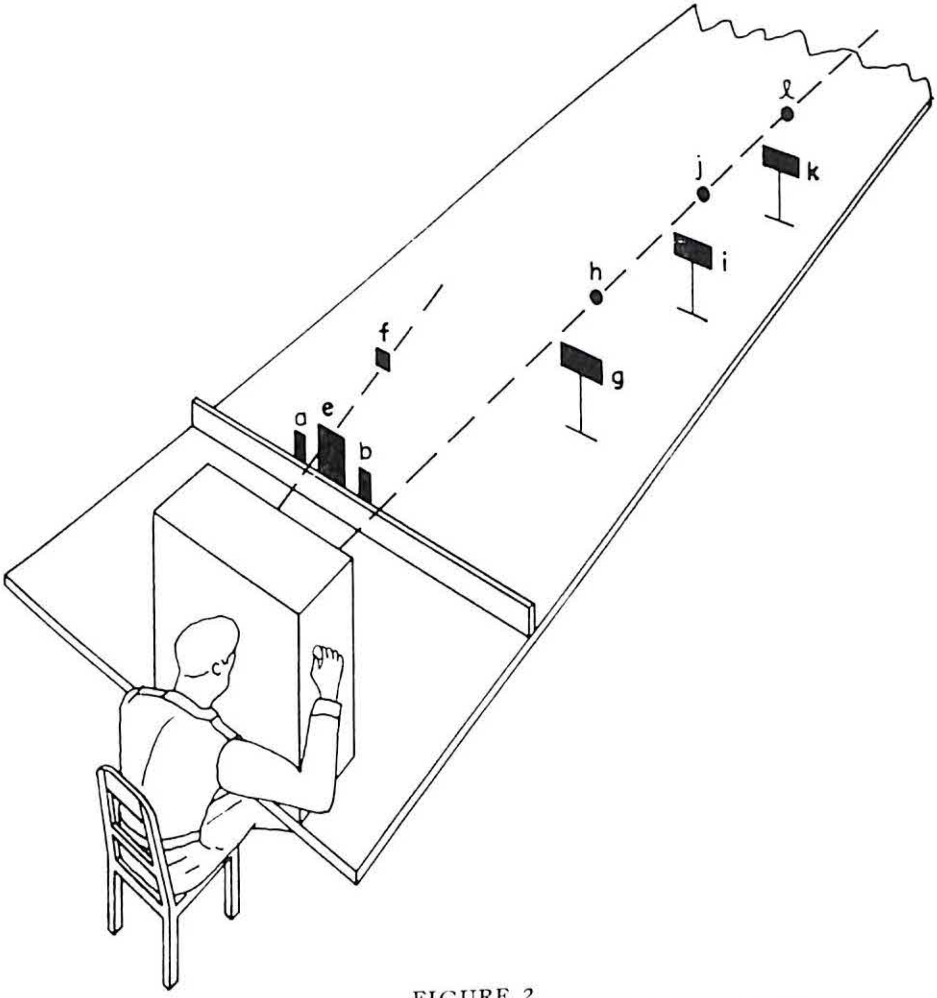

FIGURE 2
A PERSPECTIVE SCHEMATIC DIAGRAM OF THE VISUAL ALLEY

and *k* were 6, 12, 18, and 24 feet respectively from the subject. Object *f* was a square which appeared behind Object *e*. It was always placed so that the binocular disparity between Objects *e* and *f* ($\alpha_e - \alpha_f$) was 580 seconds of arc. Objects *g*, *i*, and *k* represent three alternate positions of a rectangle which was 5 inches wide by 1½ inches high. A disc which was adjustable in distance by *O* moved in depth so as to pass ½ inch above the rectangle at *g*, *i*, or *k*. Three different positions of this disc are represented by *h*, *j*, and *l*. The binocular disparity between the disc and Objects *g*, *i*, or *k* is designated as $\alpha_g - \alpha_h$, $\alpha_i - \alpha_j$, and $\alpha_k - \alpha_l$ respectively. The task of *O* was to adjust the disc behind the rectangle at *g*, *i*, or *k* until the distance between the disc and this rectangle appeared equal to the distance between Objects *e* and *f*. During the depth adjustments, Objects *a* and *b* were not present.

In the width judgement portion of the experiment, Objects *g*, *i*, or *k* were rectangles which were either 3, 5, or 7 inches wide by 1½ inches high. As in the case of the depth adjustments, only one rectangle was presented at any one time with this rectangle being placed at either 12, 18, or 24 feet from *O*. Objects *e* and *f* and the moveable disc were not present during the width judgements. Object *a* was stationary while Object *b* could be moved laterally. The task of *O* was to specify when the distance between Objects *a* and *b* appeared to be the same as the width of the rectangle at the distance D_g, D_i, or D_k.

b. Generating the Disc and Square. The only permissible depth cue within the depth intervals X_{ef}, X_{gh}, X_{ij}, or X_{kl} was that resulting from binocular disparity. For this reason, the more distant of each pair of objects (*f* or *h* of Figure 1) was stereoscopically generated rather than being a real object. The stereoscopically generated objects were the square (Object *f*) and the disc (Object *h*, *j*, or *l*) of Figure 2. Some of the principles involved in the instrument for generating these objects (reticles) are shown schematically in Figure 3.

In Figure 3 the solid lines labeled P_1, P_2, P_3, and P_4 each represent a ray of light from sources S_1, S_2, S_3, and S_4 which, after being reflected by mirrors M_1, M_2, M_3, and M_4, respectively, passed through the nodal point *N* of the eyes. The binocular fusion of the retinal stimuli resulted in the perception of a square (Object *f*) and a disc (Object *h*). The mirrors M_1, M_2, M_3, and M_4 were partially transmitting so that, in addition to the square and disc, real objects in the field of view were also visible.

Each optical system represented by a source of light and a mirror in Figure 3 was rotatable as a unit with the right and left optical systems for the same object being linked together so as to rotate equal but opposite

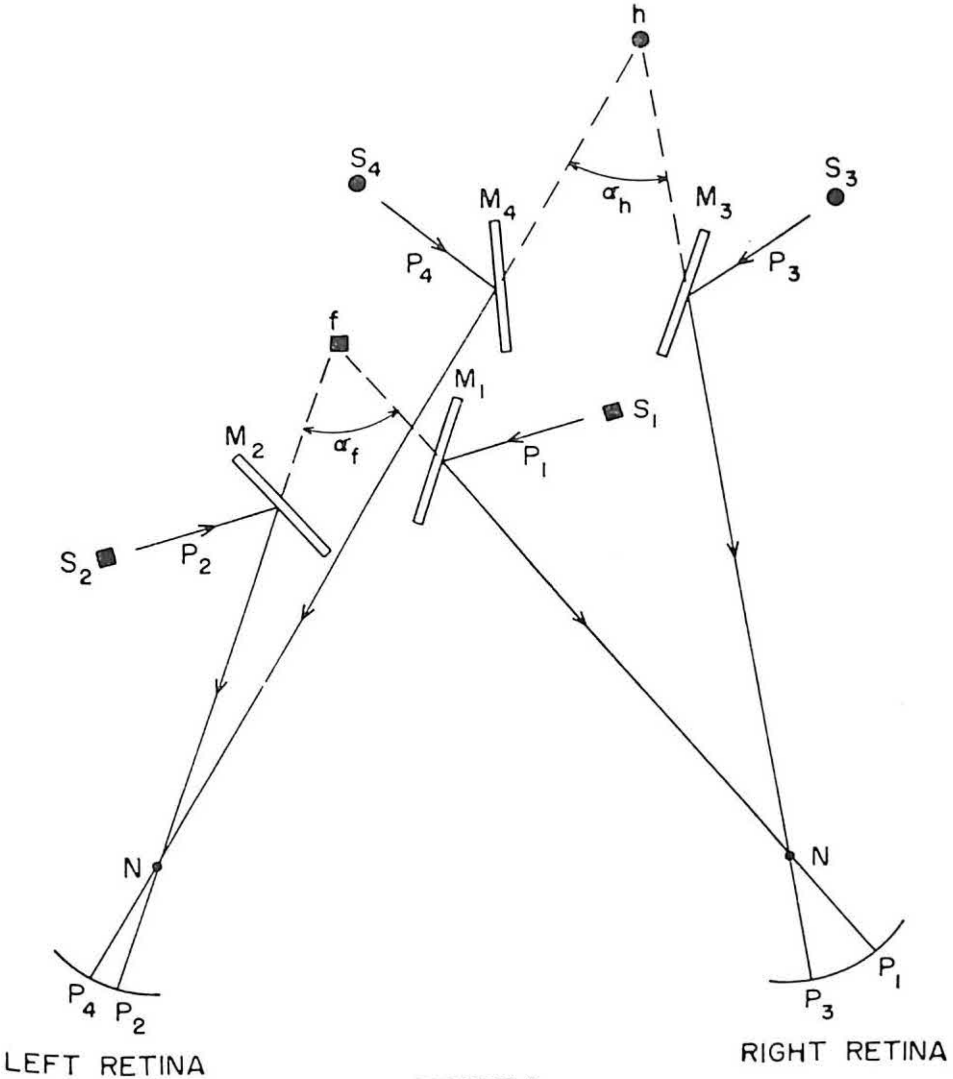

FIGURE 3
A SCHEMATIC DIAGRAM OF THE METHOD OF PRODUCING THE DISC AND SQUARE

amounts. This symmetrically opposite rotation permitted *O* to move the disc toward or away from himself in apparent depth by changing the binocular disparity between the disc and real objects in the field of view. The apparent distance position of the square (Object *f*) remained constant for an *O* throughout the experiment. The accommodation to the square and disc was always equivalent to that to a very distant object.

c. Depth Cues in the Visual Alley. The floor and walls of the alley were visible to *O*. The distance position of any of the real objects with respect to the floor of the alley was evident through their supporting structures. Thus, a variety of cues were present to determine the perceived absolute distance of Objects *a, b, e, g, i,* and *k*. But, this was not the case with respect to the stereoscopically produced objects (the disc and square). These appeared to be suspended above the floor of the alley without any physical support. Furthermore, they were always along a line of sight which was directly above the line of sight to the particular object with which each was paired. This directional isolation of the disc and square from the floor of the alley enhanced the effectiveness of the binocular disparity cue[2] in producing the perceived depth intervals (6).

2. *Procedure*

Twelve subjects participated in both the depth and width judgements. Since the subjects differed in interpupillary distance, the position was determined for each subject at which the square reticle was 580 seconds of arc behind Object *e*. The depth adjustments consisted of an adjustment of the disc by *O* to apparent equidistance with the left edge of the rectangle (*g, i,* or *k*) and, following this, an adjustment of the disc behind this rectangle so as to perceptually duplicate the perceived distance between the square reticle and Object *e*. The difference between these two adjustments, as measured by a dial indicator, specified the amount of binocular disparity at the particular distance which appeared to *O* to be equal to the standard depth interval of 580 seconds of arc. Three such pairs of adjustments were completed at each of the three distances. The order of presenting the rectangles at the three distances was systematically varied between *O*s. The *O*s also made a verbal estimate of the distance (*D*) to each of the rectangles.

For the width judgements, the experimenter, from a position invisible to *O*, moved Object *b* laterally until the distance between Objects *a* and *b* apppeared to *O* to be the same as the right-left extent (width) of the rectangle at either 12, 18, or 24 feet. Both the order of presenting the three distance positions of the rectangle and the width of the rectangle used at each distance was systematically varied between *O*s with all three sizes of rectangles being used with each *O*. Following the width judgements, *O* made verbal estimates of the distances to the rectangles in feet.

---

[2] Another but probably less appropriate method of eliminating extraneous cues to relative depth is given in a previous report (see Experiment I, Reference 3). This report also contains a more detailed account of the present experiment.

C. RESULTS

The median results from the depth and frontal adjustments are given in Table 1. Medians rather than means are presented due to a tendency for the data to be skewed by the results from one *O*. In Table 1, the physical depth interval required in order for this depth interval to remain perceptually unchanged increased with distance (*D*). The nature of this relationship is shown in Figure 4 which demonstrates that *X* is directly proportional to *D*. Conversely, Table 1 also shows that the required binocular disparity is inversely related to *D*. The last row of Table 1 gives the median results obtained from dividing the width (*S*) of the rectangle at D_g, D_i, or D_k by the adjusted width at D_e. If the adjusted width at D_e had been physically equal to the widths of the rectangles at the various distances, all the ratios in this row would have been unity. This would be described as perfect frontal size constancy. It appears that within errors of measurement, perfect frontal size constancy was present even though, from the results shown in Table 1, perfect relative depth constancy was not.

TABLE 1
MEDIAN RESULTS FROM THE DEPTH AND WIDTH JUDGMENTS

	Conditions at D_e	Results at D_g, D_i, D_k		
	$D_e = 6$ ft.	$D_g = 12$ ft.	$D_i = 18$ ft.	$D_k = 24$ ft.
Binocular Disparity	580 sec.	310 sec.	190 sec.	140 sec.
X	15 cm.	32 cm.	45 cm.	58 cm.
$\frac{S_{g.i.k}}{S_e}$	1.00	1.01	1.03	.95

The average verbal reports of distance were 4.7, 9.8, 14.1, and 17.7 feet for the 6, 12, 18, and 24 foot distances respectively. This shows the usual underestimation of absolute distance which occurs in situations of this type (5). It is evident however that the black rectangles were perceived as being definitely placed in the visual field.

D. DISCUSSION

1. *General Theory*

In a previous study (5) it was hypothesized that for small binocular depth intervals, the perceived ratio of a binocular depth to an adjacent frontal

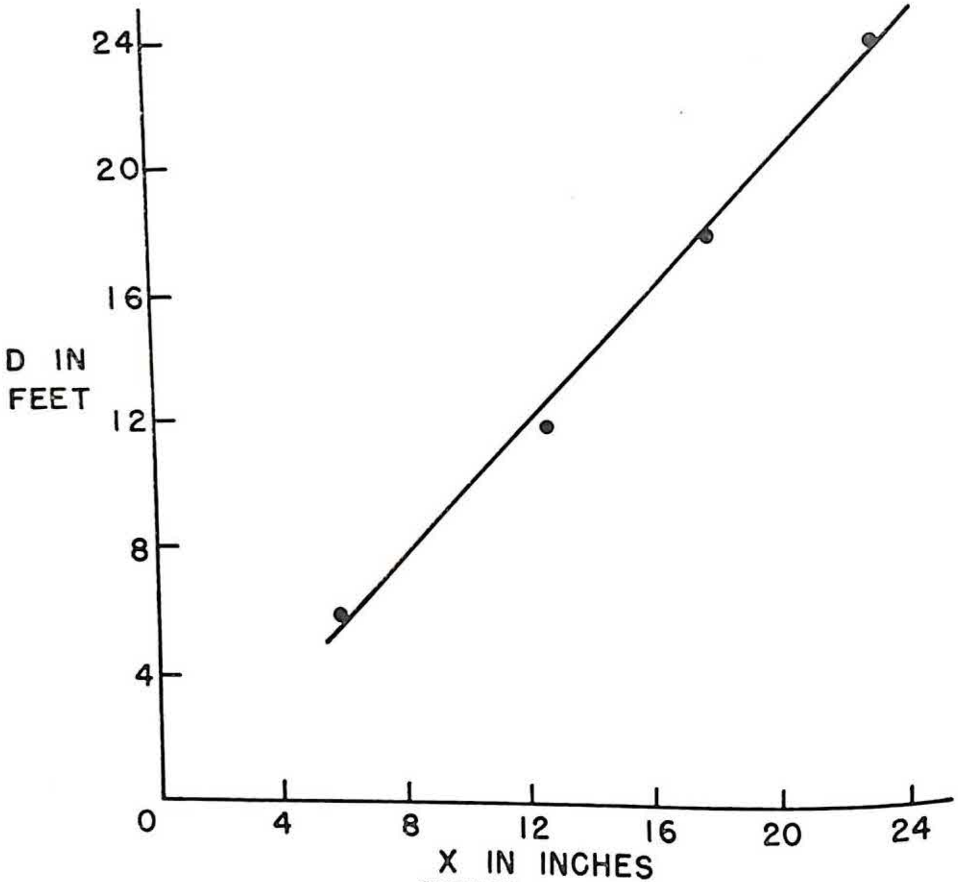

FIGURE 4
The Physical Depth Intervals (X) at Different Distances (D) Required for Perceptual Equality

width is proportional to the ratio of the angular representation of these two extents on the retina. As was discussed (5), it follows from this that

$$(\alpha_e - \alpha_f)' = \frac{1}{C}\int_f^e \frac{E_v}{\theta_v}\, d\alpha \qquad (2),$$

where,

$(\alpha_e - \alpha_f)'$ = the perceived depth associated with the binocular disparity $\alpha_e - \alpha_f$.

$\frac{E_v}{\theta_v}$ = the perceived size (E_v) per unit of angular frontal size (θ_v) associated with a frontal plane at any distance (D_v).

$d\alpha$ = a differential amount of binocular disparity.

C = an individual observer constant.

In order to transform Equation 2 to a useful form it was necessary to express $\frac{E_v}{\theta_v}$ as a function of α. For this purpose several additional relations are required. Two of these, in which α and θ are expressed in radians are given as follows:

$$\alpha = \frac{A}{D} \tag{3}$$

and

$$\theta = \frac{S}{D} \tag{4}.$$

Two additional relations are also required. Unlike Equation 2, these relations do not express a theoretical point of view. One of these relations was developed to express E_v as a function of the amount of frontal size constancy present. This equation is merely a way of noting how the perceived size of a frontal extent which is physically constant will change with distance if different amounts of frontal size constancy are present. This relation is

$$\frac{E_v}{E_e} = \frac{D_e^n}{D_v^n} \tag{5}.$$

In Equation 5, the term E_v is the perceived size of a constant *S* at any distance D_v. The term E_e is the perceived size of the same *S* at the distance D_e and *n* is a term which indicates the amount of frontal size constancy which is present. It was noted (5) that zero amount of frontal size constancy is indicated by $n = +1$, perfect frontal size constancy is indicated by $n = 0$ and, an amount of over frontal size constancy is indicated by $n = -1$. For simplicity it is assumed that for a particular set of conditions, *n* remains constant throughout the portion of the visual field being considered. Thus, *n* is a convenient way of describing how $\frac{E}{\theta}$ changes with distance. The value of *n* must be known before Equation 5 can specify how the perceived size of a constant *S* varies with distance.

A second relation to be used in conjunction with Equation 2 is like Equation 5 in that it does not express a theoretical position. It merely provides an alternative expression for *E*, the perceived size of *S*. This expression is that, the perceived size E_v of a constant physical size *S* at any distance D_v can be written as

$$E_v = K_{nv}\ S \tag{6}.$$

Equation 6 states that the perceived size of *S* at a distance can be expressed as the physical size *S* multiplied by a factor *K*, where *K*, as indicated by the subscripts *v* and *n*, may vary as a function of the distance D_v and the amount of frontal constancy *n*. For the same frontal size *S* at distances D_e and D_g, Equation 6 becomes,

$$E_e = K_{ne}\ S \qquad (7)$$

and

$$E_g = K_{ng}\ S \qquad (8).$$

The term *K* in Equations 6, 7, and 8 merely relates physical and perceived frontal size. When $n = 0$, i.e., when perfect frontal size constancy occurs, *K* is the same at all distances for a given observer. For all other values of *n*, *K* varies with distance. The reason for introducing the term *K* is to have a term which is related to perceived size and yet is not dependent upon the particular frontal size *S* which was used. For example, if at the same distance *D*, the extent *S* is doubled, the perceived size *E* of *S* will also double but within limits *K* will remain unchanged. It is easily demonstrated by combining Equations 5, 6, 7, and 8 that

$$K_{nv}\ D_v{}^n = K_{ne}\ D_e{}^n = K_{ng}\ D_g{}^n \qquad (9)$$

for all values of *D* and *n*. Combining Equations 2, 3, 4, 5, and 7 gives:

$$(\alpha_e - \alpha_f)' = \frac{K_{ne}\ D_e{}^n}{C\ A^{n-1}} \int_f^e \alpha_v{}^{n-1}\, d\alpha \qquad (10)$$

and similarly,

$$(\alpha_g - \alpha_h)' = \frac{K_{ng}\ D_g{}^n}{C\ A^{n-1}} \int_h^g \alpha_v{}^{n-1}\, d\alpha \qquad (11).$$

In the present experiment, a more distant depth interval was made perceptually equal to a nearer, standard depth interval. Therefore in the present experiment $(\alpha_e - \alpha_f)' = (\alpha_g - \alpha_h)'$. Also, from Equation 9, $K_{ne}\ D_e{}^n = K_{ng}\ D_g{}^n$. The result from applying these relations to Equations 10 and 11 is that

$$\int_f^e \alpha_v{}^{n-1}\, d\alpha = \int_h^g \alpha_v{}^{n-1}\, d\alpha \qquad (12).$$

But, according to the experimentally obtained results in the last row of Table 1, perfect frontal size constancy was present in this experiment and therefore, $n = 0$. Substituting $n = 0$ in Equation 12, using Equation 3, and simplifying, results in

$$\frac{D_f - D_e}{D_e} = \frac{D_h - D_g}{D_g} \qquad (13).$$

But, from Figure 1, $D_f - D_e = X_{ef}$ and $D_h - D_g = X_{gh}$. Therefore,

$$\frac{X_{ef}}{X_{gh}} = \frac{D_e}{D_g} \quad (14).$$

Equation 14 gives the predicted relation between X and D. This will be compared with the obtained relation shown in Figure 4.

2. *Relation Between Obtained and Predicted Results*

Equation 14 states that under conditions of perfect frontal size constancy, a perceived depth interval which is supported by binocular disparity cues will remain perceptually constant at different distances only if the physical depth interval X is directly proportional to the distance D. But, this is exactly the result which is demonstrated by Figure 4. The decrease in the obtained binocular disparity with increasing D shown in Table 1 also is in agreement with the point of view of this study as can be readily demonstrated by the appropriate combination of Equations 3 and 13. It is clear therefore that the results obtained in this study offer considerable support for the point of view expressed by Equation 2.

It will be noted that the derivation of Equation 12 from Equations 10 and 11 requires that C remains constant for the particular O at the several distances. The agreement between the predicted results from Equation 14 and the obtained results of Figure 4 offer support for the assumption that the observer constant C was unaffected by D or α. This means that from both this study and the previous one (5), there is experimental evidence for regarding C as an observer constant which can be applied unchanged to perceptions occurring at any distance.

3. *Some Relations to Other Studies*

In a previous study, the theoretical viewpoint which has been discussed in this paper has been applied successfully to the problem of perceived three dimensional shape (5). This, together with its successful application in the present experiment to the problem of the perception of a depth interval suggests that the basic equations are valid. In still another study (4) evidence was presented that the perception of a relative depth interval with binocular observation could occur independently of perceived absolute distance. As discussed in this study, it was maintained that this circumstance is logically consistent with the point of view of the present paper. It is asserted that if perceived absolute distance has any effect upon the perceived depth resulting from binocular disparity this is only a consequence of its effect on size constancy. Whether or not this effect occurs is problematical (2, 7, 8, 12). However, this is of no consequence for this study. It is only necessary to determine the amount of size constancy which is present in order to apply the above equations. It is not necessary to consider the cause of the size constancy. Thus, the above equations can be applied quite independently of the cue systems which may be present to determine perceived absolute distance as long as only binocular disparity is present to support perceived relative depth. In this respect, the present point of view can be more generally applied than that described by the Luneburg theory of binocular vision (9, 10) which has not been formulated to apply to environments in which so called monocular cues are present to contribute to the perceived position of the configuration of stimuli being judged. A discussion by Squires (14) and Shipley and Walker (11) concerning a study by Squires (13) is relevant to this problem.

It must be emphasized again, however, that the above equations do not apply to situations in which monocular cues are effective within the relative depth intervals being judged. Gilinsky had *O*s adjust a depth interval at different distances so as to be perceptually equal to a memorized foot rule or meter stick (1). The experimental situation was such that monocular cues were probably present to help determine the perceived relative intervals and thus the equations of the present study are not necessarily expected to apply. In both the study by Gilinsky and this one, relative depth intervals were increasingly underestimated with increasing distance from *O*. Evidently, in a number of situations, depth in the visual world contracts with respect to the physical world as distance increases.

E. Summary and Conclusions

This study applied a previously developed hypothesis to the problem of the perceptual duplication of a depth interval at different distances from *O* using binocular disparity cues. Both frontal size judgements and judgements of depth intervals were obtained in a situation in which there was a variety of cues for the perception of absolute distance. However, the only cue supporting the perception of the depth interval was that of binocular disparity. As expected, the amount of depth interval required at one distance to perceptually duplicate a given depth interval at another distance was directly proportional to the distances. It is asserted that this relation between physical distance and the perception of a depth interval is dependent upon an observer constant and the amount of frontal size constancy present.

REFERENCES

1. GILINSKY, A. S. Perceived size and distance in visual space. *Psych. Rev.*, 1951, **58**, 460-482.
2. ———. The relation of perceived size to perceived distance: An analysis of Gruber's data. *Amer. J. Psychol.*, 1955, **68**, 476-480.
3. GOGEL, W. C. Apparent depth duplication with binocular disparity cues. *US Army Med. Res. Lab.*, 1958, Report No. 354, 1-23.
4. ———. Perceived frontal size as a determiner of perceived binocular depth. *J. of Psychol.*, 1960, **49**, 119-131.
5. ———. The perception of shape from binocular disparity cues. *J. of Psychol.*, 1960, **50**, 179-192.
6. GOGEL, W. C., & HARKER, G. S. The effectiveness of size cues to relative distance as a function of lateral visual separation. *J. Exper. Psychol.*, 1955, **50**, 309-315.
7. GRUBER, H. E. The relation of perceived size to perceived distance. *Amer. J. Psychol.*, 1954, **67**, 411-426.
8. ———. The size-distance paradox: A reply to Gilinsky. *Amer. J. Psychol.*, 1956, **69**, 469-476.
9. HARDY, L. H., RAND, G., RITTLER, M. C., BLANK, A. A., & BOEDER, P. The geometry of binocular space perception. Report of the Knapp Laboratory of Physiological Optics, Institute of Ophthalmology, Columbia Medical Center, to ONR, New York, 1953.
10. LUNEBURG, R. K. Mathematical Analysis of Binocular Vision. Princeton, N. J.: Princeton Univ. Press, 1947.
11. SHIPLEY, T., & WALKER, G. L. Under "correspondence": Luneburg theory of visual geodesics in binocular space perception. *A.M.A. Arch. Ophth.*, 1957, **57**, 311-312.
12. SMITH, O. Judgments of size and distance in photographs. *Amer. J. Psychol.*, 1958, **61**, 529-538.
13. SQUIRES, P. C. Luneburg theory of visual geodesics in binocular space perception. *A.M.A. Arch. Ophth.*, 1956, **56**, 288-297.
14. SQUIRES, P. C. Under "correspondence": Luneburg theory of visual geodesics in binocular space perception. *A.M.A. Arch. Ophth.*, 1957, **57**, 312-314.

US Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 271-278.

THE DETERMINATION OF PERCEIVED MOVEMENT DIRECTION*

Department of Psychology, Michigan State University

MARTIN KRAMPEN AND HANS H. TOCH[1]

A. PROBLEM

The present study was designed to explore the role of directional connotation in the perception of arrows and arrow-like figures. The hypothesis we tested was the following: In an ambiguous stroboscopic (apparent movement) situation, an arrow should tend to move in the direction of its "head." Further, the more "arrow-like" a stimulus becomes, the stronger should this tendency be.

By "ambiguous stroboscopic presentation" we mean one in which a flash of light is followed or preceded by two lights flashing simultaneously, equidistant from the first light. If the two lights follow the center light in time, the perceived movement is away from the center; if the center light flashes last, movement is seen toward the center. As has been indicated elsewhere (1,2), one can "load" this type of presentation so that movement is seen predominantly to the left or right. The extent to which one is successful in making movement appear unilateral, provides a measure of one's ability to convince one's subject that the stimulus presented to him has a psychologically meaningful direction.

The present study asks the question: How much like an arrow does an arrow have to look before its "arrowness" becomes an effective determinant of perceived movement direction?

B. METHOD

1. *Apparatus*

The apparatus used was a viewing unit containing three separate compartments with two 7.5W. filament bulbs each, facing the subject with a one way transparent mirror, as sometimes used in sun glasses. Between the mirror and the bulb compartments three-partitioned transparencies, 2.5″ × 2.5″ each, could be inserted, by means of a slide holder. The bulbs in the three compart-

---

* Received in the Editorial Office on June 28, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This experiment was supported under terms of Research Grant G-8610 from the National Science Foundation entitled "Studies in apparent movement."

ments were connected to a set of five electronic interval timers wired in series, which provided accurate control over the sequence of flashes, the duration of each flash, and the interval between flashes. Brightness level was determined by means of a variable transformer connected to a voltmeter. The stimulus slides, once inserted, centered over the light compartments. In the present experiment, the two "outside" stimuli were flashed simultaneously for 500 msec. After an interval of 15 msec. the center figure was flashed for 500 msec.[2]

2. *Subjects*

Sixty-eight undergraduate students in psychology were used as subjects for the study. They were briefly acquainted with the apparatus at the start of the session. A short period of dark adaptation followed. Each subject sat seven feet from the viewing screen, with his head located on a headrest. Each stimulus was presented at least three times, and more often if requested by the subject.

3. *Procedure*

Subjects were told to fixate "approximately the center of the screen." They were further instructed as follows:

> This experiment explores some of the phenomena related to apparent movement which you know from neon signs or the movies. Watch the screen and tell me what happens. You may see figures which move or do not move. If they move, they may move from the outside to the center of the screen or mainly toward one side of the screen, either to your right or to your left. In some of the figures the movement may be stronger or faster than in others. Please report immediately as you observe it. I can repeat the trials for you as often as you want.[3]

4. *Stimuli and Design*

The experiment utilized a total of nine stimuli, which were presented in one of two series. These are represented in Figure 1. Both series contained a symmetrical control stimulus, shaped somewhat like a barbell, and a slight modification of this figure, featuring primitive tail. These two designs are those numbered 1 and 2 respectively in Figure 1. Both series also contained a "complete" or "ideal" arrow design (No. 6 in Figure 1). Three intermediate designs, with increasing "arrowness," differed in the two series. In one (consisting of Nos. 3a, 4a, and 5a) emphasis is placed on increasing differentation of the arrow's head; in the other series (3b, 4b, and 5b) the figures contain an explicitly designed tail.

---

[2] The flash duration and time interval used were those recommended by Von Schiller (2).

[3] These instructions were intended to convey as much of the universe of possible perceptual resolutions as could be comprised in them. More unstructured instructions used in preliminary experimentation had been found to yield relatively undifferentiated percepts. This finding will be discussed in another context.

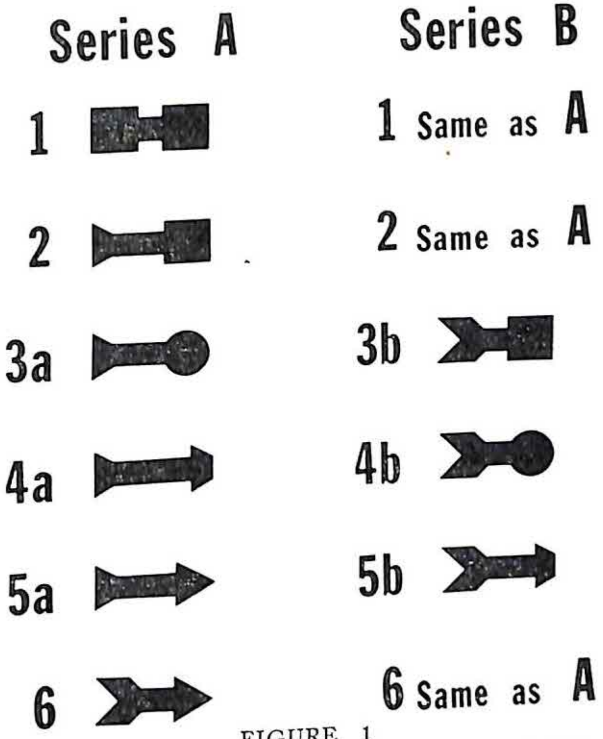

FIGURE 1

Designs Used as Stimulus Figures for the Study

The two series were presented both in "ascending" order (starting with the symmetrical figure and ending with the arrow) and "descending" order (arrow through symmetrical figure). The direction in which the figures were pointed was alternated in consecutive presentations. Half the subjects saw the first arrow design pointed to the left, and half saw it point to the right. Table 1 represents the design in schematic form, and shows the number of subjects used under each of the conditions.

C. Results and Analysis

If one discounts the symmetrical control slide, 198 out of a total of 340 presentations (43 per cent) produced movement in the predicted direction. Among subjects who faced the presentations in "descending" order (with the complete arrow presented first), the difference in favor of "correct" movement

TABLE 1
SCHEMATIC REPRESENTATION OF DESIGN, WITH NUMBER OF SUBJECTS IN EACH CONDITION (TOTAL N=68)

		Number of subjects "Ascending" Order of Presentation (from "Neutral" to "Perfect Arrow")	 "Descending" Order of Presentation (from "Perfect Arrow" to "Neutral")
Design Series A			
	Alternating sequence starting with 1st arrow pointed left	7	9
	Alternating sequence starting with 1st arrow pointed right	9	10
Design Series B			
	Alternating sequence starting with 1st arrow pointed left	9	9
	Alternating sequence starting with 1st arrow pointed right	7	8
Total Subjects		32	36

is statistically significant beyond the 0.1 level. There was also an effect of directional connotation when the figures were presented in "ascending" order (with the arrow last), but this difference is less striking ($x^2 = 9.06$, $p < .02$).

More detailed results follow. For the purpose of tabulation, the groups who saw the first presentation pointed left have been combined with those who started with the first figure pointed right. Since no significant differences were found between these groups, they were regarded as stemming from the same population.

1. The data with respect to *directional connotation* have been recorded in Table 2. It was found that in the "ascending" order of presentation, the control stimulus and the "ideal" arrow design were relatively consistently perceived as moving in the direction predicted by the hypothesis ($p < .01$). In the "ascending" series Stimulus 3b was perceived as moving toward the center significantly more often than would be expected by chance ($p < .05$). There were no indications of gradually increasing effectiveness of the various other designs.

In the "descending" sequence, the control stimulus did not show the effectiveness in directional connotation which it had shown in the "ascending" sequence. Stimulus 3b, on the other hand, showed highly significant directional

TABLE 2

Direction of Movements Perceived with Various Arrow Designs: Movement in The Direction "Pointed to," Movement in Opposite Direction, And Movement Toward the Center (N=68)

		"Ascending" order of presentation			"Descending" order of presentation		
		Movement perceived in					
Stimulus #		"Correct" direction	"Incorrect" direction	Center	"Correct" direction	"Incorrect" direction	Center
Symmetrical control	(1)	28***	4		15	21	
	2	10	7	15	12	11	13
	3a	7	5	4	5	8	6
	3b	2	4	10*	13**	2	2
	4a	9	3	4	12**	1	6
	4b	5	4	7	11*	3	3
	5a	6	5	5	5	7	7
	5b	7	2	7	6	3	8
Perfect arrow design	(6)	20**	6	6	18*	5	13

*** Difference significant beyond .01 level (x^2, 1dβ).
** Difference significant beyond .01 level (x^2, 2dβ).
* Difference significant beyond .05 level (x^2, 2dβ).

connotation. Stimulus 4a and Stimulus 4b proved significant in their directional connotation. Stimulus 6 (the "ideal" arrow design) also moved in the direction predicted by the hypothesis to a significant degree ($p<.05$). When the designs of "gradually increasing arrowness" were tested for effectiveness, a definite directional connotation was found at work in the intermediate stimuli 3a, 3b, 4a, 4b, 5a, 5b combined ($p < .01$), in the "descending" sequence. On the whole, the "descending" order of presentation, and the stimuli with the more pronounced tail design, seemed to produce more percepts in the direction predicted by the hypothesis:

2. *Order of presentation* is clearly important. For example, Stimulus 3b was perceived in the "ascending" order of presentations to move predominantly toward the center (from both sides), whereas this same stimulus in the "descending" sequence was perceived moving *in the direction of its head* to a highly significant degree.

Statistical tests of the influence of the perception of stimuli on the perception of next following stimuli only in two cases (5a vs. 4a, and 4a vs. 3a in the descending sequence) showed a significant carry-over effect, ($p<.02$ in both cases). We shall return to this observation below.

3. We also tested the *difference in effectiveness of the two series of designs,*

Series *A* (with undifferentiated tails) and Series *B* (with differentiated tails). In the "descending" order of presentations, Stimulus 3b (with pronounced tail) proved to be significantly more effective than Stimulus 3a. Other tests suggest that there is a difference between Stimulus Series *A* and *B* in favor of *B,* but these differences fail to reach a satisfactory level of significance. The differences in effect of design appeared to have been reduced by other factors.

4. There seemed to be a definitive difference in effectiveness between some of the *stimuli presented in "ascending" order and the same stimuli presented in "descending" order*. This difference was found to be highly significant for the control stimulus and Stimulus 3b (in both cases beyond the .01 level). The differences point to a set factor at work: When an "ascending" sequence starts with the control stimulus it is likely that the subject perceives (correctly) movement to the center, because he has no false expectancies formed by previous presentations. On the other hand, when a "descending" sequence terminates with the presentation of the control stimulus, the subject brings with him directional expectancies formed through the perception of the previous stimuli which all have some degree of directional connotation. He thus tends to see the symmetrical stimulus moving in one direction or the other rather than toward the center. In the "descending" sequence, subjects saw the complete arrow moving to the center more often than when the arrow was presented at the end of the "ascending" sequence after the subject had had cumulative experience of directionality. With respect to the difference between 3b "ascending," and 3b "descending," the following interpretation seems to emerge from the data: The start of the "ascending" sequence with a neutral stimulus did not build up any directional expectancies in the subject, and Stimulus 2, with weak directional connotation, was not in fact perceived as moving significantly more often in one direction or the other. This double experience of non-directionality could build up in the subject an expectancy (set) for non-directionality such that *Stimulus 3b would now be expected to move to the center rather than in either direction.* With increasing "arrowness" in the design this set dissipates as the "ideal" arrow is approached. Conversely, the "descending" sequence can build up through the initial presentation of the "ideal" arrow a strong directional expectancy, which grows cumulatively with each presentation, reaching a maximum at Stimulus 3b.

The fact that neither Stimuli 5a nor 5b were effective, although they represent a design with a high degree of "arrowness" can be explained by invoking an additional expectancy. The "ideal" arrow could set the stage not only for directional expectancy in general, *but also for directional expectancy in the specific direction in which it pointed*. Since the directions of the stimuli were alternated, the subject, relying on the expectancy built up by the arrow, had to perceive Stimulus 5 as moving in the "wrong" direction. The perception of the "wrong" direction of Stimulus 5, in turn sets the stage for the "correct" perception of Stimulus 4. This interaction process is sustained and enhanced by the cumulative set of directional expectation in general.

D. Conclusion

We are able to confirm our hypothesis, but within limitations. We did find our complete arrow design to be more effective in determining movement direction than the various incomplete or imperfect designs we tried. And it does appear as if some types of arrow designs (those possessing tails) are more effective than others. But no fine gradations of effectiveness materialized.

There is no question but that several of our more striking results are a function of expectancies occasioned by order of presentation. Two types of set appear to be involved: One (most obviously present in the results obtained with our "descending" sequence) may be characterized as the expectation that movement will occur *in the direction in which the previous presentation was perceived to move*. The second type of set involves the expectation of *unilateral or bilateral movement*. The subject approaches a given presentation as if saying to himself "I assume, from my immediate experience, that these presentations move toward the center," or "it appears that these lights may be expected to move toward one side or the other."

The situation is further complicated by the instructions which, it will be recalled, specified a variety of possible percepts. This provides, in a sense, a generalized set, which probably facilitates the formation of particular sets, as well as sensitizing the subject to the possibility of fairly differentiated percepts (irrespective of their determinants). The analogy which comes to mind is that of a hypochondriac who by virtue of his condition easily locates the slightest pain, but also finds his malady difficult to abandon when the physical occasion for it is no longer present.

A final word may be in order concerning our presentation and the special nature of the percepts it elicits: Our presentation is not really "ambiguous." The sequence of lights calls for equilateral movement inward, toward the center. This movement, however, is much weaker than a unilateral one, in that it contains two component movements, which are subject to accentuation. What we attempt to do, in experiments such as that reported above, is to accentuate by causing the perceptual process *to take into account characteristics other than the sequence of lights.* These characteristics then enter as over-determinants; if strong enough, they supercede the determinant of sequence.

The process may be regarded as one of relative weighting, and as such, *it permits the assignment of weights.* What becomes possible is the design of experiments in which various effects of past experience can be pitted against each other. With optimal designs, hierarchies of perceptual determinants may be teased out of the interaction products.

REFERENCES

1. TOCH, H., & ITTELSON, W. H. The role of past experience in apparent movement: A revaluation. *Brit. J. Psychol.*, 1956, **47**, 195-207.
2. VON SCHILLER, P. Stroboskopische Alternativversuche. *Psychol. Forsch.*, 1933, **17**, 180-214.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 279-285.

RELATIONSHIPS BETWEEN EXTRAVERSION, NEUROTICISM, AND ASCENDANCE*[1]

Payne Whitney Psychiatric Clinic, New York Hospital

A. BRONZAFT, R. HAYES, L. WELCH, AND M. KOLTUV

A. PROBLEM

In the field of personality testing there has been much interest in correlations among traits. Studies concerned with extraversion, neuroticism, and ascendance often yield confusing and conflicting results and the question arises to what extent these personality traits have been clearly defined and are mutually exclusive.

A questionnaire called the Maudsley Personality Inventory for the measurement of extraversion and neuroticism has been recently developed by Eysenck (7). Eysenck (7) reports that for subjects who have taken the Maudsley, there is no significant correlation between extraversion and neuroticism. He uses this claim to support his theory that these traits are independent and unrelated. Jensen (9), in discussing Eysenck's data on the Maudsley Personality Inventory, states that in the general population only a slight correlation exists between introversion and neuroticism. He suggests that even this slight relationship may be partly "an artifact due to a factor such as differences in the 'social desirability' of the introverted and extraverted items" (9, p. 322).

However, Bernreuter (4), Hoitshma (8), Moore and Steele (10) have found high correlations between introversion and neuroticism. Bernreuter (4) has gone so far as to state that neurotic tendency correlates with introversion and these "are names given to a single trait whose real nature has been obscured by inadequacies of tests by which it has been estimated" (4, p. 402). Conklin (6) contends that the tests of extraversion-introversion have been measuring abnormal manifestations of these traits and this may account for some of the confusion in the area.

Several investigators (3, 5), using the Allport A-S scale and either Laird's Personal Inventory or the Heidbreder test of introversion-extraversion, have found a low but positive relationship between ascendance and extraversion.

---

* Received in the Editorial Office on July 17, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] From the Department of Psychiatry, Cornell University Medical College (Payne Whitney Psychiatric Clinic) and the Department of Psychology, Hunter College.

The Allports, in discussing the low correlation, say that it "indicates that ascendance and submission are distinctive and should not be loosely identified with other variables in personality" (1, p. 16).

Stagner (11), using the Allport A-S scale and the Thurstone Personality Schedule, obtained a correlation of —.51 between ascendance and neuroticism. Allport (2), using the same two personality tests, obtained a correlation of .35 between submission and neuroticism. Some of the correlation Allport attributed to the 11 identical items in both tests, for when these were removed the correlation shrunk to .25. Allport (2) concludes that there is a tendency for the neurotic to be submissive but it is not marked enough to justify identity of the concepts of neuroticism and submission. However, other investigators (10) have found a close relationship between submission and neuroticism.

The following hypotheses were tested in the studies which follow:

Hypothesis 1: Extraversion and neuroticism are negatively correlated.

Hypothesis 2: Extraversion and ascendance are positively correlated.

Hypothesis 3: Ascendance and neuroticism are negatively correlated.

B. PROCEDURE

The subjects were 46 female college students between the ages of 17 and 32 years, enrolled in psychology courses at a large municipal college. Since participation was a course requirement, they were not volunteers.

The Maudsley Personality Inventory developed by Eysenck was used to measure introversion-extraversion and neuroticism. The Allport A-S Reaction Study was used to measure ascendance-submission. The two tests were administered to the subjects at different times.

C. RESULTS

Means and SD's were computed for the subjects' scores on the A-S scale and on the extraversion and neuroticism scales of the Maudsley Personality Inventory. Table 1 gives the means and the SD's for all the variables.

Table 2 presents the correlations between the variables. Only the correlations for extraversion and neuroticism and extraversion and ascendance were

TABLE 1

MEANS AND SDs ON THREE VARIABLES (N = 46)

Variables	Mean	SD
Extraversion (MPI)	27.63	8.06
Neuroticism (MPI)	20.02	10.28
Ascendance (Allport)	1.22	17.98

TABLE 2
Relationships Between Extraversion, Neuroticism, and Ascendance (N = 46)

Variables	Correlations
Extraversion-neuroticism	—.346*
Extraversion-ascendance	.548**
Neuroticism-ascendance	—.223

* Significant .05 level.
** Significant .01 level.

significant. It is noted that there is a significant inverse relationship between extraversion and neuroticism at the .05 level. The positive correlation of .55 between extraversion and ascendance is significant beyond the .01 level. The correlation of —.22 between ascendance and neuroticism was not significant, but in the same direction as those reported by Stagner (11) and Allport (2).

D. Discussion

Jensen, in presenting Eysenck's data on the Maudsley Personality Inventory, states that a significant negative correlation is obtained between extraversion and neuroticism "only in those samples which in some way represent some highly selected (and therefore biased) element of the general population" (9, p. 322). He cites as an example a group of prisoners. However, his table of normative data (Table 3) also shows a significant correlation at the .05 level between extraversion and neuroticism for 200 normal adult males, a significant correlation at the .01 level for 145 American university students of both sexes, and another significant correlation at the .01 level for a combined group of 22 female student nurses and 64 mixed English university students.

A comparison of the means and SD's on the extraversion and neuroticism scales for our group and for a group of 350 American female university students cited by Jensen (9) reveals striking similarity. This is also true for the 145 mixed American university students and both our group of American female college students and the group of mixed American university students show significant negative correlations for the two traits. Jensen (9) gives no correlation between extraversion and neuroticism for the female university students.

From the above evidence one might conclude that extraversion and neuroticism are negatively correlated. However, Eysenck in discussing the relationship between extraversion and neuroticism states that "the independence of the two scales was demonstrated by the low correlation of —.09 for the original sample of 400 men and women, and the even lower correlation of —.07 for

TABLE 3

STANDARD DEVIATION, RELIABILITY, AND INTERCORRELATION OF THE EXTRAVERSION AND NEUROTICISM SCALES OF THE MPI FOR VARIOUS GROUPS (9)

		E-Scale		N-Scale			Reliability	
Sample	N	Mean	SD	Mean	SD	r_{EN}	E	N
1. Normal adult males	200	24.62	10.04	17.81	11.32	—.15*	.85^{a}	.90^{a}
2. Normal adult females	200	25.17	9.33	19.45	11.02	—.04	.82^{a}	.87^{a}
3. Total of 1 & 2	400	24.89	9.67	18.63	11.19	—.09	.83^{a}	.88^{a}
4. English university students, male	50	28.86	8.36	19.04	11.24	.12		
5. English university students, male	213	25.26	8.85	23.23	11.27	—.07		
6. English university students, mixed	64	25.16	10.22	26.78	9.28			
7. Student nurses, female	22	23.82	9.71	30.64	9.22			
8. Total of 6 & 7	86	24.81	10.11	27.77	9.42	—.30**		
9. Polytechnic & art school students, mixed	68	24.57	19.04	27.06	11.56	—.08		
10. American university students, male	714	28.40	8.06	20.19	10.71			
11. American university students, female	350	29.41	8.37	21.63	10.45			
12. American university students, mixed	145	27.77	7.60	21.57	9.75	—.20**	.74^{b}	.84^{b}
13. Industrial apprentices, male	100	29.34	9.00	21.20	10.26	—.09		
14. Industrial apprentices, male	76	29.18	8.19	21.21	10.41	—.14		
15. Neurotic patients, male	83	19.09	10.33	32.98	10.78			
16. Neurotic patients, female	65	18.67	9.21	34.75	11.83			
17. Total of 15 & 16	148	18.91	9.86	33.75	11.29	—.30**		
18. Dysthymics^{c}	25	21.00	11.96	36.80	10.48			
19. Hysterics & psychopaths	27	25.22	9.96	28.82	12.76			
20. Recidivists, male	72	24.76	10.08	32.18	10.38	—.32**		

* = .05.

** = .01.

[a] Corrected split-half reliability.

[b] Kuder-Richardson "Formula 20".

[c] "Dysthmic" is Eysenck's term for neurosis of the anxiety, depression, obsessive-compulsive, and phobic types.

a further male group of 200" (7, p. 14). However, in addition to presenting these two correlations Jensen (9) cites several significant correlations between extraversion and neuroticism. These correlations contradict Eysenck's statements concerning the independence of the extraversion and neuroticism scales.

Jensen (9), in his paper in support of the Maudsley, is forced to admit that some of the correlations cited in his article are not in accord with Eysenck's theory. Jensen (9) sees no explanation for this fact and suggests that Eysenck's theory may not be correct in all particulars. However, Jensen still concluded his paper by saying, ". . . the MPI may be regarded as the preferred measure of introversion-extraversion and neuroticism" (9, p. 324). He goes on to say that although it needs further improvements, ". . . the MPI in its present form can be recommended for research purposes as being perhaps the best questionnaire measure of introversion-extraversion and neuroticism available at the present time" (9, p. 324). From the data Jensen (9) presents on the Maudsley, and from the results we obtained, it is difficult to accept his conclusions. Even if we had never done our experiment but worked with Jensen's data only, we would have the same arguments.

Since there are definite differences between the American and English cultures, there is some question concerning the suitability of the Maudsley Personality Inventory for Americans. The Maudsley in its present form with English norms may be yielding this significant correlation between introversion and neuroticism in American groups due to the cultural differences. Whereas Americans are encouraged to speak out, Englishmen are reticent in their behavior. Thus, if the Maudsley is to be suitable with Americans, we may have to use different norms.

We have no information except in the specified cases whether the populations are English or American or of some other national background. In addition we don't know whether the subjects were volunteers or not. Different populations might yield different results, but since our mean and SD were not very different from the ones reported by Jensen (9), we do not expect this to be the case generally.

Although earlier investigators report a low but positive correlation between ascendance and extraversion, our correlation of .55, significant beyond the .01 level, could not be termed low. This may very likely be due to the fact that we are using a different test to measure extraversion-introversion. Both the Maudsley and the Allport A-S scale may be measuring the same thing to a large extent and this may account for the high correlation between the two. When we compared our mean score to the one the Allports (1) reported, our mean, while within the average range, was closer to the submissive end of the scale. Our group was unselected in that they were not volunteers and the submissives who may ordinarily shy away from such experiments were present in our sample. These two factors may account for the different findings.

Although some studies report a close relationship between submission and neuroticism, our correlation in the same direction was not significant. It is probably noteworthy that we used the Maudsley for the test of neuroticism whereas other investigators used the Thurstone Personality Schedule. There are only 23 items measuring neuroticism on the Maudsley scale and several of these appear to be redundant. Thus, the Maudsley may only be measuring some aspects of neuroticism which may account for the difference between our results and previously reported ones.

The correlation between neuroticism and introversion was significant but the correlation between neuroticism and submission was not. Although it was stated previously that our measures of extraversion-introversion and ascendance-submission may be measuring the same thing to a large extent, the concepts of extraversion-introversion and ascendance-submission are not identical.

E. Summary and Conclusions

The Maudsley Personality Inventory and the Allport A-S Reaction Study were administered to 46 female college students. A positive correlation was hypothesized between extraversion and ascendance while a negative correlation was hypothesized between extraversion and neuroticism, as well as between ascendance and neuroticism. Eysenck (7), who developed the Maudsley, reported that the extraversion and neuroticism scales are independent. However, our results are contradictory to his, confirming our hypothesis that extraversion and neuroticism are negatively correlated. The high positive correlation obtained between extraversion and ascendance is also contradictory to the statement of the Allports that only a low positive correlation exists between the two. Although earlier investigators have reported a negative correlation between ascendance and neuroticism, our correlation in the same direction was not significant. Discussion is centered about the differences between previous findings and our findings.

References

1. Allport, G. W., & Allport, F. H. The A-S Reaction Study: Manual of Directions, Scoring Values, and Norms. (Second Revised Edition). New York: Houghton Mifflin, 1939.

2. Allport, G. W. Neurotic personality and traits of self-expression. *J. Soc. Psychol.*, 1930, **1**, 524-527.
3. Bender, I. E. Ascendance-submission in relation to certain other factors in personality. *J. Abn. & Soc. Psychol.*, 1928, **23**, 137-143.
4. Bernreuter, R. G. The theory and construction of the personality inventory. *J. Soc. Psychol.*, 1933, **4**, 387-405.
5. Broom, M. E. A study of a test of ascendance-submission. *J. Appl. Psychol.*, 1930, **14**, 405-413.
6. Conklin, E. S. The determination of normal extravert-introvert interest differences. *J. Genet. Psychol.*, 1927, **34**, 28-37.
7. Eysenck, H. J. A short questionnaire for the measurement of two dimensions of personality. *J. Appl. Psychol.*, 1958, **42**, 14-17.
8. Hoitsihma, R. K. Reliability and relationships of the Colgate Mental Hygiene Test. *J. Appl. Psychol.*, 1925, **9**, 293-303.
9. Jensen, A. R. The Maudsley Personality Inventory. *Acta Psychol.*, 1958, **14**, 314-325.
10. Moore, H., & Steele, I. Personality test. *J. Abn. & Soc. Psychol.*, 1934, **29**, 45-52.
11. Stagner, R. The intercorrelation of some standardized personality tests. *J. Appl. Psychol.*, 1932, **16**, 453-464.

Payne Whitney Psychiatric Clinic
New York Hospital
525 East 68 Street
New York 21, New York

Published as a separate and in *The Journal of Psychology*, 1960, 50, 287-290.

A FURTHER STUDY OF WEAK-LIGHT REINFORCEMENT AND RESPONSE FACILITATION*

Department of Psychology, University of Mississippi

FRANKLIN A. NASH, JR. AND WILLIAM F. CROWDER[1,2]

A. INTRODUCTION

In a study recently reported in this journal (1), weak-light onset was found to possess little power to reinforce bar pressing in rats. That experiment employed a two-bar Skinner box with presses on one of the bars bringing the light. It was therefore thought that the task of learning to press one of a pair of identical levers might have been too "difficult" to provide a sensitive test of the light's reinforcing properties. Accordingly, the present experiment substituted a single-bar box for the two-bar box used previously.

B. METHOD

1. *Subjects and Apparatus*

The *S*s were 58 Holtzman male albino rats, about 75 days old. They were housed two per cage, with water and food always available. The apparatus consisted of a pair of Skinner boxes made from wire-mesh living cages, 9 in. wide, 14 in. long, and 7 in. high. Each box contained a sheet steel lever which was 3-¾ in. wide and 5 in. above the floor. A 6-w., 115-v. clear glass lamp was mounted 6 in. above the bar. A Fiberglas-lined cover shielded the box from sound and light.

2. *Procedure*

The experiment lasted two days: one day of lever training and one day of testing for retention of the lever training effects. The *S*s were run in pairs. A randomly selected member of each pair was assigned to the experimental group. The other *S*s were assigned approximately equally and at random to the two control groups: "conventional" and "yoked-box."

a. Lever training. On this day the *S*s were placed in Skinner boxes for 25 min. The first 10 min. was an habituation period in the dark. During the remaining 15 min., each lever press by an experimental S turned on a 1-sec. light for it, and, if its control S was a yoked-box animal, for the latter as well. Thus any yoked-box control S received exactly the same distribution of light flashes as did a particular experimental S. The rest of the control Ss ("conventional" controls) did not receive the light at any time.

---

* Received in the Editorial Office on July 18, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to thank Drs. J. B. Morris, W. A. Wilbanks, and J. B. Wolfe for their critical reading of this paper.

[2] This study was supported in part by grants from the Committee on Faculty Research and the Annual Alumni Giving Program.

b. Retention test. On the day following lever training, each S was placed in a Skinner box for 20 min. and the number of lever presses was recorded. The light was not presented.

C. Results and Discussion

Table 1 shows the mean number of lever presses for each group during lever training and also during the retention test. Table 2 shows the results of t tests performed on differences between groups on each day.

TABLE 1

Mean Number of Lever Presses Made on Training Day and on Retention Test Day

Group	N	Training	Retention
Experimental	29	65.2	43.2
Yoked-box control	14	36.6	38.5
Conventional control	15	27.8	39.6
Both control groups	29	32.0	39.1

TABLE 2

Significance of Differences Between Experimentals, Yoked-box Controls, Conventional Controls, and Combined Controls

	t	df	p
Training Day			
Experimentals vs. yoked-box controls	4.05	41	.001
Experimentals vs. conventional controls	5.66	42	.001
Yoked-box vs. conventional controls	2.06	27	.05
Retention Test Day			
Experimentals vs. yoked-box controls	.56	41	.50
Experimentals vs. conventional controls	.48	42	.60
Experimentals vs. combined controls	.69	56	.50
Yoked-box vs. conventional controls	.17	27	.80

Three indexes of reinforcement are shown. The first one is obtained by comparing the experimental group with the "conventional" or no-light control group *during* lever training, as most previous studies have done. The experimental Ss pressed the bar more than twice as frequently as did the conventional controls (65.2 vs. 27.8 responses). This difference was highly reliable according to the t test ($p < .001$), confirming the findings of previous investigations. The second index is based on the comparison of the experimental group with the yoked-box control group, again during lever training. Significantly more responses were made by the experimental *S*s than by yoked-box controls (65.2 vs. 36.6 responses; $p < .001$). Hence, by this more stringent test, the light still appears to have been reinforcing. A similar finding is reported by Kling, *et al.* (3), who also used yoked-box controls. The third index ("retention") compares the experimental group with both control groups on the day following lever training. The experimental group pressed slightly more often than did either control group (43.2 vs. 39.6 and 38.5 presses respectively). Neither difference approached statistical significance: both levels of confidence exceeded .50. The retention scores of all 29 control *S*s were then combined and were compared with those of the 29 experimental *S*s. Again, the experimental-control difference failed to reach significance ($p = .50$). Thus what had seemed by the first two indexes to be a marked reinforcing effect during the lever training could not be detected on the retention test. Had the experimental *S*s *learned* to press the bar, they should have retained the "habit" for at least a day. The results of the present experiment suggest that the light did not reinforce bar-pressing, but only served to facilitate it.

On the other hand, several studies have obtained clear evidence for light reinforcement while using retention measures. Striking results are reported by Forgays and Levin (2), who employed a two-lever apparatus with presses on one lever bringing a 5-sec. light. When the reinforced lever was reversed after ten days of training, the *S*s continued to press the previously reinforced bar, even though the other bar was being reinforced on that day. These and other findings attest to the genuineness of the light reinforcement phenomenon.

The conclusions that do seem warranted by the results of the present experiment and the previous one (1) are these:

1. Light reinforcement does not appear under all conditions where it might be expected. The crucial conditions necessary for light to act as a reinforcer have not been identified, however.
2. Light onset can *appear* to reinforce even when it does not actually do so.
3. By inference, any experimental treatment that affects the light's power to evoke bar-pressing will also appear to influence the latter's reinforcing power, unless suitable controls for evocation are employed.
4. The "yoked-box" method is not an adequate control for response facilitation or evocation.

D. SUMMARY

The purpose of this study was to investigate the reinforcing power of the onset of a weak light, while controlling for any direct facilitating or evoking effects of the light. Rats were given 15 min. of lever training in a Skinner box, with each response producing a 1-sec. presentation of light. Just as in previous studies, the rats receiving the light for responding pressed more often than did rats not receiving the light. This difference could have been due to reinforcement, facilitation, or both. A partial control for facilitation was provided by presenting the light to a third group of *S*s ("yoked-box" controls) whenever the experimental *S*s responded and received the light. Again, more responses were made by the experimental *S*s. In an attempt to eliminate all possible facilitation effects, an additional test for reinforcement was made. This test consisted of placing *S* back in the Skinner box on the day following lever training, and recording the bar-presses but never presenting the light. It was supposed that the light's reinforcing effects—if any—would be retained for at least one day, but that the light, being absent, could not elicit or facilitate the lever responses. The combined control groups (N = 29) made almost as many responses as did the 29 experimental animals. The difference between the groups reached only the 0.50 level of significance. These data do not appear to support the hypothesis that the onset of a weak light acts as a reinforcer with white rats.

REFERENCES

1. CROWDER, W. F., WILKES, W. P., & CROWDER, T. H. Weak-light reinforcement with and without control for response facilitation. *J. of Psychol.*, 1960, **49**, 181-184.
2. FORGAYS, D. G., & LEVIN, H. Discrimination and reversal learning as a function of change in sensory stimulation. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 191-194.
3. KLING, J. W., HOROWITZ, L., & DELHAGEN, J. E. Light as a positive reinforcer for rat responding. *Psychol. Rep.*, 1956, **2**, 337-340.

Department of Psychology
University of Mississippi
University, Mississippi

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 291-301.

SIMILARITY AS A FACTOR IN RECALL: I. THE EFFECT OF COLOR*

Department of Psychology, University of Arizona

DOROTHY IRENE MARQUART

A. INTRODUCTION

The theories of the nature of the effect of similarity upon recall seem to fit into several basic groups. One group of theorists maintains that elements or patterns of the present situation (*C*) remind the subject of the stimulus that was present during the original learning (*c*) which in turn produces the response (*d*). According to Höffding (8), who originated this theory, *C* brings forth *c* because of the operation of a "selective effect" of similarity. Höffding said that this effect is independent of the laws of learning. Ward (30) and Köhler (14, 16) agree with Höffding's theory of the order of events in recall (from *C* to *c* to *d*) but differ in their ideas of the mechanisms responsible. Ward said that the effects of similarity can be explained by the law of contiguity. And Köhler extended the theory to include the ideas that similarity is a pattern effect, that similarity effects in recall are similar or identical to similarity effects in perception, and that it is a recognition process.

The second group of theorists maintains that, since the two stimuli have the same neural locus, there is no need to assume that *C* recalls *c*. Each of these theorists said, also, that similarity effects are dependent upon some one of the basic laws of learning. This type of view has been favored by Thorndike (28, 29), William James (12), and Hollingworth (9).

A third theory was presented by Householder and Landahl (10) and by Landahl (17). These authors maintain that the effects of similarity are dependent upon the operation of two independent mechanisms: similarity and difference. They say that similarity effects are directly proportional to the number of neurones that are excited in common by two stimulus patterns. If two stimuli differ in any way, even in intensity, spatial position or temporal position, they excite different groups of neurones in the brain. If the difference in cortical neurones excited is considerably larger than that required to perceive a difference in stimuli, recall will not occur. That is, similarity effects will be lacking.

---

* Received in the Editorial Office on July 22, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

E. S. Robinson (25) objected to all three theoretical views discussed above. He insisted that we can never know the neural elements involved, thus we should not discuss them. His law of assimilation says that similarity is important in recall but makes no attempt to explain the fact.

Despite the amount of theorizing that has occurred, very little experimental evidence is available. Köhler and von Restorff (16) and Bartel (2) have demonstrated that the important factor in spontaneous recall situations is distinctive similarity rather than similarity as such. According to these experiments, similarity serves as a factor in recall only if interpolated events are dissimilar to those which are common to the two critical groups of tasks. They demonstrated this by showing that interpolated problems which are similar to original problems but which do not allow the applications of the principles learned during the solution of the original problems cause these principles to be carried over to the final problem less frequently than do interpolated dissimilar problems.

Köhler (14, 16) has attributed this phenomenon to the fact that in the favorable situation *c* and *C* belong to the same class but the interpolated materials do not. He said that if *c* and *C* have much in common with the interpolated material " . . . *C* (and probably also the trace of *c*) will tend to become indifferent members of the large and quasi-homogeneous sequence, and the specific 'step backward' from *C* to *c* will be less likely to occur" (14, p. 138).

Shepard and Fogelsonger (26) performed an experiment which Köhler (15) contends shows the importance in recall of similarity of pattern. Shepard and Fogelsonger had their subjects learn pairs of nonsense syllables some of which had identical second members. They found that the recall of the second syllable was better when only one association had been formed and only one syllable was given as a cue than when two had been formed and both cue syllables were given. Since, in the second situation mentioned above, two cues were present for each recall, the results were unexpected. Köhler said that the presentation of two syllables as a cue would make a situation look unfamiliar to the subject.

Studies which have been conducted to determine the effects of the similarity of the learning and recall environments and of the learning and recall responses have yielded diverse results. Wilson, in an unpublished study described by Smith and Guthrie (27), found that nonsense syllables were relearned more readily in the room in which the original learning occurred than in a different room, and that the presence of identical odors on the two occasions facilitated relearning. Reed (23) failed to find any significant differences produced by similar or dissimilar postures, sense modalities, and response forms.

One of the difficulities in the study of these phenomena has been in the establishment of a criterion of similarity. Landahl (10, 17) said that similarity can be measured by the number of elements common to the two stimulus situations. Duncker (5) maintained that it is a gestalt phenomenon and that we must talk of similarity of pattern rather than of elements, and Pepper (21) said that similarity must be defined in terms of opinions rather than in terms of common elements or patterns. He said that two things are similar if they are viewed as similar.

Most of the experiments which have been performed to determine whether gestalten or elements are the true units of similarity have found patterns to be the more important (4, 13, 18, 20). Attneave (1), however, claims that deviations from similarity are best measured by a sum of differences rather than by a pattern of differences.

The method of measurement of similarity used in the studies mentioned above has involved generalization. The difficulty with this method stems from the existence of mediated generalization which depends upon learning. It is true that Maltzman and Brooks (19) deny that semantic generalization, a form of mediated generalization, has been demonstrated. However, among others, Foley, Cofer and their co-workers (3, 6), Hull (11), Razran (22), and Reiss (24) have reported its existence. If generalization is to be used as a measure of similarity, it is necessary to go as far as Hilgard and Marquis (9) and accept the concept of mediated similarity.

The present series of experiments was conducted for the purpose of determining the nature of the effects of similarity upon learning and recall. Will paired nonsense syllables which are reproduced in a single color be learned more rapidly than paired nonsense syllables reproduced in different colors? Will the color used for the cue syllable influence the number of nonsense syllables recalled? The answers to these questions should supply evidence of importance for the selection of a theory of recall.

In order to accomplish this purpose it was necessary to: (*a*) Ask the subjects to learn paired nonsense syllables. (One-half of the pairs had the syllables printed in the same color. The other half of the pairs had the two syllables printed in different colors.) (*b*) Conduct an immediate test using the first syllable of each pair as printed on the learning cards as a cue syllable, and (*c*) Conduct a second test using cue syllables printed in the same color as during the original learning or differing from the original syllable in color. (In Experiment II the cue syllable for this test was, on occasion, the same color as the recall syllable but not as the cue syllable.)

B. Experiment I

1. *Apparatus and Subjects*

The apparatus used in this study consisted of two sets of 12 white cardboard cards (11" by 3½") and of four sets of 12 smaller cards (5½" by 3½"). Six of each set of the larger cards contained a pair of nonsense syllables printed in a single color. That is, both members of the pair were one color but no two cards were printed in the same color. The other six cards had two syllables printed upon them, each a different color. The use of two sets, rather than one set of cards, was to control the effect of differences in the learning difficulty of the pairs. The pairs of syllables used were the same for the two sets of cards, but those pairs printed in one color on one set were printed in different colors on the other set.

Each of the smaller cards had the first member of one of the pairs found on the larger cards printed upon it. One-half of these cards had this syllable printed in the same color as on the larger card, and one-half were printed in colors which had formerly not been used in the experiment. The six cards which contained a syllable printed in the same color as the first syllable on the corresponding larger card were distributed so that three of them contained the first member of a pair which was printed in the same color on the larger card and three of them contained the first member of a pair which was printed in different colors. The same division was made for those printed in colors not formerly used in the experiment.

Thus there were four groups of card-pairs each consisting of a large card and a small card (see Table 1). (*a*) There were three pairs with the two syllables printed in the same color on the larger card and with the syllable on the small card printed in this color. We shall call these the same-same pairs. (*b*) There were three pairs with the two syllables printed in the same color on the larger card and with the printing on

TABLE 1
THE FOUR TYPES OF CARD-PAIRS

Card pair designation	Larger cards		Small cards	
	Same color	Different colors	Same color as first syll.	Formerly unused color
Same-same	X		X	
Same-different	X			X
Different-same		X	X	
Different-different		X		X

the small card a second color. We shall call these the same-different pairs. (*c*) There were three pairs with the two syllables printed in different colors on the larger card and with the printing on the small card the same color as the first syllable on the larger card. We shall call these the different-same pairs. And (*d*) there were three cards with the two syllables printed in different colors on the larger card and with the syllables on the small card printed in a third color. We shall refer to these as the different-different pairs.

The use of four sets of small cards was to control the effect of possible differences in the learning difficulty of the various pairs.

One-hundred seventy subjects were used in this experiment. All were students enrolled in an intermediate level psychology course at the University of Arizona. They were studied in groups of from 8 to 20 each. All of the subjects were naive regarding the problem being studied although all had had considerable experience in learning nonsense syllables.

2. *Procedure*

A set of the larger cards was first presented to the subjects five times using an exposure rate of approximately five seconds per card. Each presentation was made in a different predetermined order to prevent serial learning. The subjects were directed to learn the syllables so that they could recall the second syllable when the first was presented. Immediately after the fifth presentation, a test was given in which each large card was exposed with the second syllable covered. Approximately 10 seconds were allowed for the recall of each syllable.

The subjects were then asked to hand the materials to the instructor and were given a lecture on course material. Without warning and 20 minutes after the conclusion of the last learning trial, one of the sets of smaller cards was presented with directions requiring the subjects to reproduce the second member of the pair.

One-half of the subjects had exposed to them for learning one of the sets of larger cards, one-half the other set of larger cards. One-fourth of the subjects (one-half of each of the groups above) saw each of the four sets of small cards.

3. *Results*

a. The effects of similarity of color during learning. Immediately after learning, the first syllable on the larger card was exposed and the subject was to recall the second. More syllables were recalled when the two syllables were the same color during learning than when they were different colors.

Of those printed in the same color during learning, an average of 2.94 syllables were recalled in comparison with an average of 2.59 of those printed in different colors during learning. This difference attains a critical ratio of the difference of correlated means of 3.20.

Twenty minutes after learning 2.51 syllables were recalled which were the same color during learning, and 1.91 were recalled which were different colors during learning. The critical ratio of the difference of correlated means is 5.77.

b. The effect of the color of the cue syllable. When one compares the results of the four subgroups on the second test to determine whether the color of the cue syllable is important, one finds that when the two syllables were originally presented in a single color, it helped to have the cue syllable the same color. When the two syllables were printed in different colors, it helped to have the cue syllable a third color.

There were three pairs of syllables which were printed in the same color during learning and which had the cue syllable for the second test printed in this color (the same-same group). An average of 1.38 of these were recalled. Of the three pairs of syllables originally printed in a single color and with the cue syllable for the second test printed in a second color (the same-different group), only 1.12 were recalled. The critical ratio of the difference of correlated means is 3.33.

The average number of syllables recalled from the learning pairs which were printed in different colors with the cue syllable the same color as the first member of the pair (the different-same group) was .84. Of the three pairs of syllables originally printed in different colors and the cue syllable a third color (different-different) an average of 1.07 syllables were recalled. The critical ratio of this difference is 2.94.

C. EXPERIMENT II

1. *Apparatus, Subjects, and Procedure*

The cards used for this experiment were very similar to those used in Experiment I. The only differences were in the nonsense syllables used and in the construction of the small cards for the different-same pairs. Each of these small cards contained the first syllable from a larger card printed in the same color as the second syllable on the larger card. Different-same in this experiment means that the two syllables were printed in different colors during learning and that the cue syllable for the second test was the same color as the second member of the pair.

The subjects employed in this experiment were 147 junior, senior, and graduate students of the University of Arizona. They were obtained from the same course as those used in Experiment I. No subject was employed in both Experiments I and II.

The procedure employed was the same as that used in Experiment I. The directions were the same, the time intervals were the same, and the use of a number of sets of cards to control differences in difficulty of item-pairs was the same.

a. The effect of similarity of color during learning. Both immediately after learning and 20 minutes after learning, recall was better for those syllables which were the same color during learning than for those which were different colors during learning. Table 2 shows the averages and the critical ratios of the differences in correlated means.

TABLE 2
THE NUMBER OF SYLLABLES RECALLED WHICH WERE PRINTED IN THE SAME COLOR AND IN DIFFERENT COLORS DURING LEARNING (EXPERIMENT II)

Test	Mean same	Mean different	Critical ratio of the difference
First	3.93	2.99	7.85
Second	2.99	2.56	5.24

b. The effect of the color of the cue syllable. The difference between the second recall scores of the same-same and the same-different pairs is smaller than that in Experiment I but is in the same direction. The recall average for the same-same pairs is 1.56, that for the same-different pairs is 1.38. This difference has a critical ratio of only 2.15.

Of the syllables which were different colors during learning, more syllables were recalled on the second test when the cue syllable was the same color as the second syllable of the pair than when it was a third color. The average number recalled of the different-same group is 1.44, of the different-different group is 1.15. The critical ratio of the difference is 3.41.

D. DISCUSSION

The results of these two studies show that paired nonsense syllables which are printed in one color are learned better during a limited number of trials and recalled better after 20 minutes than are paired nonsense syllables which are printed in different colors. Also, when the two syllables are originally printed in the same color, recall is greater when the cue syllable is this color than when it is a second color. When the two syllables were originally printed in different colors, it helped to have the cue syllable the same color as the second syllable (Experiment II) but was detrimental to have it the same color as the first syllable (Experiment I).

The two most commonly accepted theories of the effect of similarity upon recall are the theory originally presented by Höffding and favored by Ward and Köhler, and the theory that similar stimuli affect identical parts of the cortex. According to both of these theories, printing the cue syllable the same color as the first syllable of the pair should facilitate recall. Printing the cue syllable the same color as the second syllable of the pair should not increase recall.

Adding Köhler's concept of a pattern effect to Höffding's theory makes it possible to explain the results of Experiment I but not those of Experiment II. It is possible to postulate greater similarity in pattern between different-different pairs than different-same pairs.[1] Such a relationship, if it does exist, would explain the greater recall of syllables from the different-different pairs than from the different-same pairs in Experiment I. However, it would lead one to predict the same type of results from Experiment II; a prediction which was not fulfilled.

Landahl's theory postulating the existence of separate similarity and difference mechanisms may help but not in the way originally presented. As Landahl stated his theory (10, 17), similarity depends upon identical neural excitation. A cue syllable which is the same color as the first member of the learning pair should excite an area of the cortex which is closer to identical to that excited by the original stimulus than should a cue syllable printed in a different color. Therefore, the different-same pairs in Experiment I should be recalled more readily than the different-different pairs. Also, printing the cue syllable the same color as the second syllable of the learning pair should not facilitate recall because the theory involves a connection from *C* to *c* to *d*.

If one instead assumes a recall of *d* directly from *C* and an inhibitory effect from false cues as well as a facilitation effect from similarity, the results of the two experiments reported in this paper can be explained. The inhibitory effect must be regarded as operating only when something that could be regarded by the subjects as a cue is misleading.

There would be a facilitation effect for the same-same pairs of both experiments since the color of the cue syllable, being the same as the second syllable, would tend to aid in the recall of the second syllable. The same-different pairs would be recalled less well than the same-same pairs because of lack of facilitation effect in the same-different situation. The color of the cue syllable was different than that of any syllable formerly experienced thus could not serve as a cue and would not inhibit the recall of the second syllable. For Experiment I the different-same pairs would be expected to be recalled less well than the different-different pairs because of an inhibitory effect as compared with a lack of effect. The recall of the different-same pairs would be inhibited because of the fact that the color of the cue syllable had been experienced during the original learning yet could not directly bring forth the response syllable. Recall of the different-different pairs would neither be facitated nor inhibited by the color of the cue syllable because the syllable color had not been experienced formerly during the experiment. The effect upon recall of printing the cue syllable the same color as the second syllable of the pair should be one of facilitation.

---

[1] The subjects did not feel that this similarity relationship existed. However, introspective analysis may not be adequate in this situation.

An analysis of the data in terms of facilitation, neutral, or inhibitory effect yields significant differences in averages. The same-same pairs in both experiments and the different-same pairs from Experiment II would be expected to show a facilitation effect. The average number of syllables recalled from each group of three of these pairs is 1.46. The same-different and the different-different pairs from both experiments would show neither a facilitation nor and inhibition effect. The average number of syllables recalled from each group of three of these pairs is 1.17. An inhibitory effect would be expected from the different-same pairs of Experiment I. The average number recalled of this group of three pairs is .84.

This type of explanation will also explain the results of the Köhler and von Restorff (16) and the Bartel (2) studies.

E. SUMMARY

Two experiments were conducted. The subjects were first shown 12 pairs of nonsense syllables five times for learning. One-half of the pairs had the two syllables printed in the same colors, one-half of the pairs had the two syllables printed in different colors. The subjects were tested twice; once immediately after learning and once 20 minutes after learning. The second test was conducted in Experiment I using cue syllables the same color as the first member of the pair or a color not used formerly in the experiment. The second test was conducted in Experiment II using cue syllables the same color as the second member of the pair or a color not used formerly in the experiment.

It was found that paired nonsense syllables are learned better and recalled better if both members of the pair are printed the same color than if they are printed different colors. When the two syllables were printed during learning the same color, it helped to have the cue syllable printed in that color. When the two syllables were printed in different colors, it helped recall to print the cue syllable the same color as the second syllable of the pair, but hindered to print it the same color as the first member of the pair. These findings can be explained by theorizing that recall is an immediate process and that it depends upon the balance of facilitation and inhibition effects.

References

1. Attneave, F. Dimensions of similarity. *Amer. J. Psychol.*, 1950, **63**, 516-556.
2. Bartel, H. Über die Abhängigkeit spontaner Reproduktionen von Feldbedingungen. *Psychol. Forsch.*, 1937, **22**, 1-25.
3. Cofer, C. N., Janis, M. G., & Rowell, W. M. Mediated generalization and the interpretation of verbal behavior: III. Experimental study of antonym gradients. *J. Exp. Psychol.*, 1943, **32**, 266-269.
4. Cutujean, F. C. The role of identical stimulus elements in judgments of similarity. *Amer. J. Psychol.*, 1942, **55**, 18-32.
5. Duncker, K. A. A qualitative (experimental and theoretical) study of productive thinking (solving of comprehensible problems). *Ped. Sem.*, 1926, **33**, 642-708.
6. Foley, J. P., & Cofer, C. N. Mediated generalization and the interpretation of verbal behavior: II. Experimental study of certain homophone and synonym gradients. *J. Exp. Psychol.*, 1943, **32**, 169-175.
7. Hilgard, E. R., & Marquis, D. G. Conditioning and Learning. New York: Appleton-Century, 1940.
8. Höffding, H. Psychologie. Leipzig: Bendixen, 1887.
9. Hollingworth, H. L. Psychology: Its Facts and Principles. New York: Appleton. 1928.
10. Householder, N. S., & Landahl, H. D. Mathematical Biophysics of the Central Nervous System. Bloomington, Ind.: Principia, 1945.
11. Hull, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9-30.
12. James, W. The Principles of Psychology. New York: Holt, 1910.
13. Klüver, H. Behavior Mechanisms in Monkeys. Chicago: Univ. Chicago Presss, 1933.
14. Köhler, W. Dynamics in Psychology. New York: Liveright, 1940.
15. ———. Gestalt Psychology. New York: Liveright, 1947.
16. Köhler, W., & von Restorff, H. Analyse von Vorgängen im Sperenfeld, II zur Theorie der Reproduktion. *Psychol. Forsch.*, 1935, **21**, 56-112.
17. Landahl, H. D. Neural mechanisms for the concepts of differences and similarity. *Bull. Math-Biophys.*, 1945, **7**, 83-88.
18. Leeper, R., & Leeper, D. O. An experimental study of equivalent stimulation in human learning. *J. Gen. Psychol.*, 1932, **6**, 344-375.
19. Maltzman, I., & Brooks, L. O. A failure to find second-order semantic generalization. *J. Exp. Psychol.*, 1956, **51**, 413-417.
20. McKinney, F. Quantitative and qualitative essential elements of transfer. *J. Exp. Psychol.*, 1933, **16**, 854-864.
21. Pepper, S. C. A convergence theory of similarity. *Phil. Rev.*, 1937, **46**, 596-608.

22. Razran, G. H. S. Semantic, syntactic, and phonetographic generalization of verbal conditioning. *Psychol. Bull.*, 1939, **36**, 578.

23. Reed, H. J. The influence of a change of conditions upon the amount recalled. *J. Exp. Psychol.*, 1931, **14**, 632-649.

24. Reiss, B. F. Semantic conditioning involving the galvanic skin reflex. *J. Exp. Psychol.*, 1940, **26**, 238-240.

25. Robinson, E. S. Association Theory Today. New York: Century, 1932.

26. Shepard, J. F., & Fogelsonger, H. M. Studies in association and inhibition. *Psychol. Rev.*, 1913, **20**, 290-311.

27. Smith, S., & Guthrie, E. R. General Psychology in Terms of Behavior. New York: Appleton, 1921.

28. Thorndike, E. L. The Psychology of Learning. New York: Columbia Univ. Press, 1913.

29. ———. Selected Writings from a Connectionist's Psychology. New York: Appleton-Century-Crofts, 1949.

30. Ward, J. Psychological Principles. Cambridge: Univ. Press, 1918.

Department of Psychology
University of Arizona
Tucson, Arizona

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 303-313.

AN EVALUATION OF THE MAGNITUDE ESTIMATION TECHNIQUE*

Psychology Division, US Army Medical Research Laboratory

GLENN R. HAWKES

A. INTRODUCTION

The introduction of techniques for the direct scaling of subjective magnitude of sensation by Stevens has led to the formulation of a new psychophysical law (9), and, perhaps inevitably, to criticism (3). There can be no doubt that the power law is capable of expressing the form of data derived by direct estimation from a number of modalities (10). These data, in log-log plot, are fitted rather well by straight lines, with of course occasional exception (cf. 1). Further, relationships determined by this method have been shown to be useful for obtaining data necessary for equal subjective spacing of stimuli (5).

Data obtained by the method of magnitude estimation have been reported to be little affected by variations in experimental technique. Stevens (8) has shown that the position of the standard in the stimulus series, or even the absence of a standard, has little effect on the slope of a line fitted to median estimates in a log-log plot. There is, however, a dearth of explicit predictions of the circumstances under which the slope exponent should or should not change. Insofar as such predictions can be made for electrical cutaneous magnitude estimates, they are tested herein for the variables of set, area of stimulation, duration of stimulation, density of innervation, and the momentary sensitivity of the receptor (s).

B. EXPERIMENT 1: INSTRUCTIONS TO THE *O*S

Both Stevens (11) and the present writer (5) have found the slope of a line fitted to estimates of intensity for electrical stimulation of the finger to be 3.5. As the current applied to the skin is increased there is an increase in the subjective intensity of the vibratory sensation. In addition, there is also an increase in the apparent area (or extent) of the sensation. Does the slope of 3.5 represent judgments of relative intensity of area or some combination of these two aspects of the sensation? This may be inferred by instructing *O*s to estimate only intensity or area.

---

* Received in the Editorial Office on July 26, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

1. *Apparatus and Procedure*

The apparatus used in the experiments reported herein was the same as that of a previous report (5), except that a Ling-Altec Model 260A amplifier was substituted for the McIntosh MC-60. Current was applied to the pad of the index finger of the *O* through a circular electrode of 12 mm. diameter at a frequency of 1500 cps. An inactive electrode of 25 mm. diameter rested on the palm of the *O*. Duration of stimulation was 1000 msec.

Ten *O*s made estimates of intensity in one session and area in another. Half estimated area in the first session; half estimated intensity first. Ten stimulus intensities were presented (in a random order) for judgment in each session, ranging from 110 to 200 per cent of the *RL* current value in steps of 10. The standard was 130 per cent of *RL* current, and was assigned a value of "10." The standard was presented twice prior to the stimulus series.

2. *Results*

Median estimates for the judgments of intensity are shown in Figure 1 and for area in Figure 2. The slopes of the lines fitted to these data (Method of Least Squares) are 3.7 for intensity and 2.3 for area. The significance of this difference was evaluated by *t*-test (for correlated samples), using as scores the slope constants of lines fitted to the estimates of each *O*, a total of 10 pairs of slopes.[1] The *t* value of 4.16 was significant beyond the .01 level confidence.

3. *Discussion*

The significantly different slopes obtained when the *O*s are instructed to report area or intensity indicate that the method of magnitude estimation can be successfully used to obtain data on two related yet discriminable aspects of sensations elicited by the identical physical stimuli. The smaller slope constant obtained for the area judgments agrees with the subjective reports that there is a more gradual increase in area than intensity as current intensity is increased. It is interesting that the slope exponent of 3.7 obtained for the intensity estimates of the present study, where *O*s were instructed to disregard area, is quite similar to the 3.5 of previous experiments (5, 11). *O*s were apparently not confused by the absence of such instructions and may be assumed to have been making estimates of sensation intensity.

---

[1] The author is indebted for this suggestion to W. F. Battig. Personal communication of April 5, 1960.

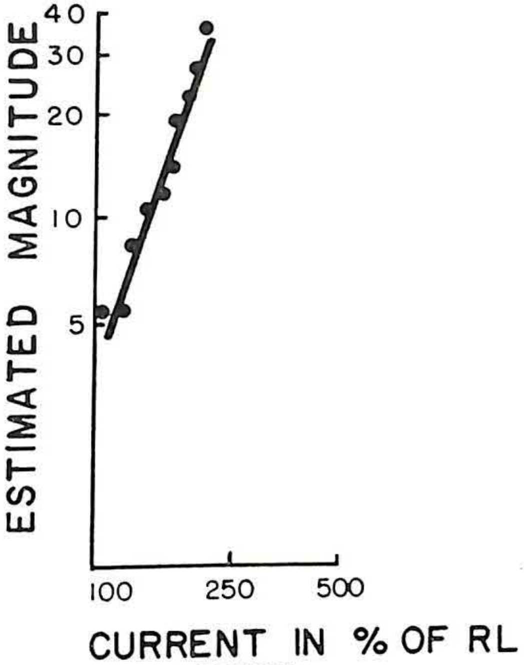

FIGURE 1
PERCEIVED INTENSITY ESTIMATES, EXPERIMENT 1

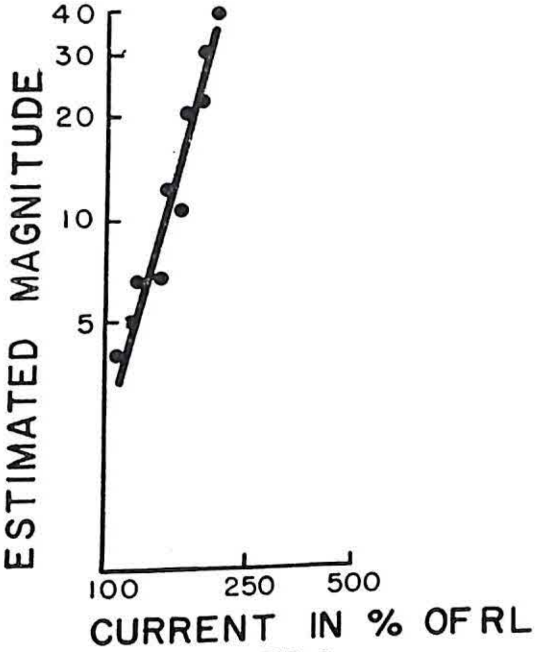

FIGURE 2
PERCEIVED AREA ESTIMATES, EXPERIMENT 1

C. EXPERIMENT 2: DURATION OF STIMULATION

Subjective reports indicate that there is a reduction in the rate of increase in apparent intensity of vibratory sensation if the stimulus duration is shortened much below 500 msec. This experiment was undertaken to determine whether or not the method of magnitude estimation would show a reduction in slope of a line fitted to such data corresponding to the reduction in subjective intensity for brief stimulus durations.

1. *Apparatus and Procedure*

The apparatus and stimulus current values were the same as used in Exp. 1. Ten *O*s made estimates of subjective intensity. Stimulus duration was 100 msec., as opposed to the 1000 msec. duration of Exp. 1.

2. *Results*

Median estimates are plotted in Figure 6. The slope of the straight line fitted to these data was 2.2. A *t*-test indicated that these results were significantly different from the intensity estimates of Exp. 1 ($t = 4.0$, $p < .01$).

3. *Discussion*

The method of magnitude estimation reflects a reduction in subjective intensity as duration of stimulation is reduced, as predicted. It follows, therefore, that one cannot specify *the* (single) exponent for electrical cutaneous magnitude estimates, for duration of stimulation is an important determinant of the slope constant.

D. EXPERIMENT 3: AREA OF STIMULATION

At a recent symposium devoted to a discussion of cutaneous research (6), there was general agreement that the stimulus for sensation intensity when electrical current is applied to the skin is current density (ma./cm.^{2}) at the receptor (s). If this is true, then the slope of 3.7 obtained for intensity estimates in Exp. 1 should not change if the area of stimulation is reduced, provided that current density and locus remain constant. Current density, however, may not affect perceived area of sensation in the same manner as perceived intensity. This experiment, then, was to determine the effect of reduction of area of stimulation on estimates of apparent intensity or apparent area.

1. *Apparatus and Procedure*

Current was applied to the pad of the index finger through an electrode of 4 mm. diameter; the inactive electrode of 8 mm. diameter rested on the palm of the *O*.

Ten *O*s made estimates of intensity in one session and area in another. Half estimated intensity in the first session; half estimated area first. Stimulus values (percentage of *RL*) were the same as in Exp. 1.

2. *Results*

Median estimates for intensity are shown in Figure 3 and for area in Figure 4. Slopes for straight lines fitted to these data were 3.1 for intensity and 2.4 for area. The t value of 1.51 indicated that area and intensity were not significantly different ($p < .10 > .05$). The intensity estimates of this study were not significantly different from the results of Exp. 1.

3. *Discussion*

The reduction in area of stimulation tended to reduce the slope for intensity estimates, but failed to affect the slope for apparent area. Further, there was no significant difference in the slopes for area vs. intensity in this study, in contrast to the significant difference obtained in Exp. 1. These data suggest that current density may have a differential effect upon apparent area vs. apparent intensity of sensation.

The failure to find a significant change in the slope exponent for intensity estimates (compared to Exp. 1) further suggests that current density is the important parameter for electrical cutaneous magnitude estimates of intensity. If a significant difference between the intensity slopes for Exp. 1 vs. Exp. 3 had been obtained, two explanations could have been offered, either that current density was not the important stimulus parameter in this study, or that current density was correct but the relationship postulated by Stevens was wrong. By the same line of reasoning, Stevens' law may hold only if the appropriate stimulus parameter has been identified and is being measured.

E. EXPERIMENT 4: DENSITY OF INNERVATION

Von Békésy (2) has shown that when density of innervation is reduced, e. g., by stimulating the forearm as opposed to the finger tip, there is a faster increase in subjective intensity for mechanical vibration of the skin as stimulus intensity is increased, and that there is a reduction of the

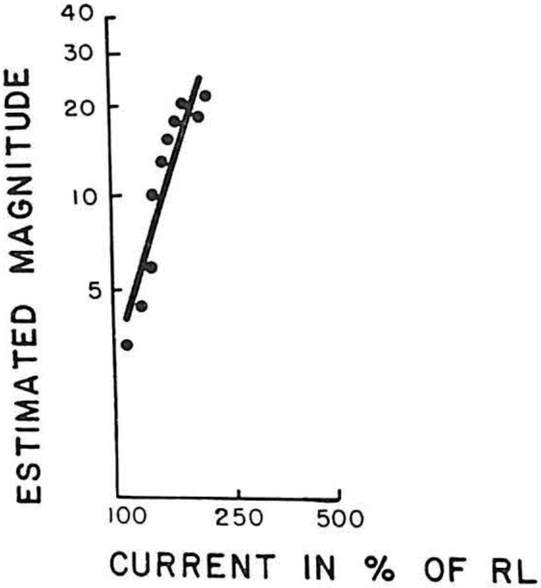

FIGURE 3
PERCEIVED INTENSITY ESTIMATES, EXPERIMENT 3

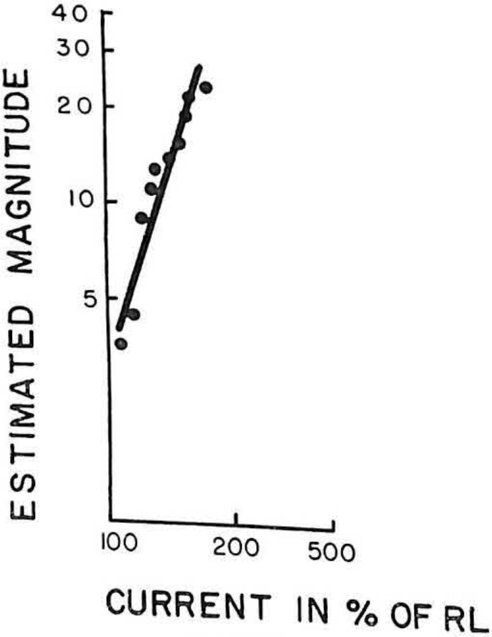

FIGURE 4
PERCEIVED AREA ESTIMATES, EXPERIMENT 3

sensory dynamic range. A faster rate of increase in subjective intensity should result in an increase in slope over the 3.7 obtained on the finger tip (Exp. 1) if the method of magnitude estimation is sensitive to variations in density of innervation.

1. *Apparatus and Procedure*

The electrodes were the same as those used in Exp. 1. The active electrode rested on the volar side of the forearm, 1 in. proximal to the wrist; the inactive electrode also was applied to the volar forearm, further toward the elbow. Ten *O*s made estimates of intensity with the stimulus values (percentage of *RL*) used in Exp. 1.

2. *Results*

Median estimates are plotted in Figure 5. The slope of the straight line fitted to these data was 3.2. This slope was not significantly different from the estimates of intensity reported in Exp. 1. A greater amount of current was necessary to reach the *RL* on the forearm than for stimulation of the finger tip.

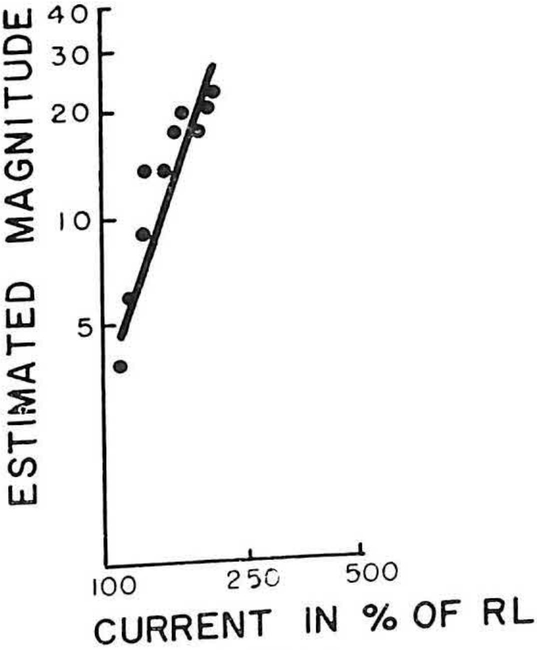

FIGURE 5
Perceived Intensity Estimates, Experiment 4

3. *Discussion*

The method of magnitude estimation failed to reflect a faster increase in subjective intensity (when an area having a reduced density of innervation was stimulated) by an increase in the slope of a line fitted to these data. One may conclude that electrical cutaneous magnitude estimation data are not affected by changes in density of innervation.

A possible explanation is that the failure to find an increase in the slope exponent may be due to the mode of operation of the receptors of the forearm. Sensations elicited by all suprathreshold current levels seemed subjectively stronger than when the finger tip was the site of stimulation. Thus, perhaps a situation occurs analogous to that of the ear when demonstrating recruitment, i. e., greater energy is necessary to reach the *RL*, but there is an increase in rate of loudness change with increase of physical intensity. More current was required to reach the *RL* (for all *O*s) on the forearm than on the finger tip. There was, however, no increase in the slope exponent for intensity estimates, so the analogy to recruitment in the ear is incomplete.

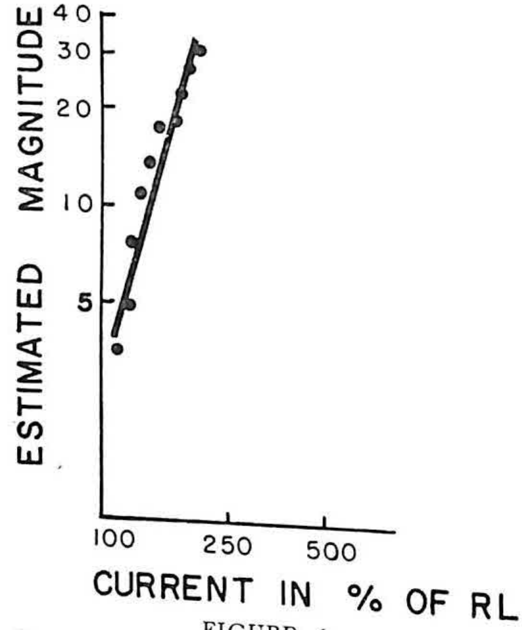

FIGURE 6
Perceived Intensity Estimates, Experiment 2

F. Experiment 5: Momentary Sensitivity of the Receptors

Warming or cooling the skin has been shown to reduce sensitivity to mechanical vibration (12), shortening the sensory dynamic range and producing a faster rate of increase in subjective intensity (2), similar to the situation in forearm vs. finger tip stimulation. This experiment was undertaken to determine whether or not the method of magnitude estimation would reflect a faster increase in subjective intensity of sensations elicited by electrical current when the finger tip was warmed. A straight line fitted to median estimates of intensity should have a slope greater than the 3.1 of Exp. 3.

1. *Apparatus and Procedure*

The apparatus was the same as that of Exp. 3. A heating unit was used to raise the temperature of the index finger. Enough heat was applied to just avoid the pain *RL*. Temperature at the finger varied from 41° to 43° C. The small electrode was used in order to facilitate heating of the finger tip.

2. *Results*

Median estimates for the ten *O*s are shown in Figure 7. The slope of a line fitted to these data was 2.9. A *t*-test indicated that the results were not significantly different from the intensity estimates of Exp. 3. More current was necessary to reach the *RL* for two of 10 *O*s as a result of warming the skin.

3. *Discussion*

Although losses of sensitivity to mechanical vibration of the finger (2) or the forearm (12) have been reported to result from warming the skin, this is not necessarily the case for electrical cutaneous stimulation. The author[2] has found such losses to occur (4), but was unable to obtain consistent losses of sensitivity for electrical stimulation when the ear lobe was warmed or cooled. It is to be noted that most of the *O*s of this study did not demonstrate losses of sensitivity (by an increase in the current required to reach the *RL*) produced by warming the finger. Any change of the slope exponent for the intensity estimates in the present experiment therefore probably was due to the masking effect of the thermal sensations.

The failure to find sensitivity losses or an increase in the slope of a line fitted to median estimates of intensity in this experiment is felt to be consistent with the view that electrical current applied to the skin first normally stimulates neural mechanisms different from those acted upon by the adequate stimulus, mechanical vibration. It has been suggested that electrical current stimulates the nerves directly, as well as the cutaneous receptors (7).

---

[2] The study referred to resulted from a suggestion by F. A. Geldard, who did extensive exploratory work on this problem.

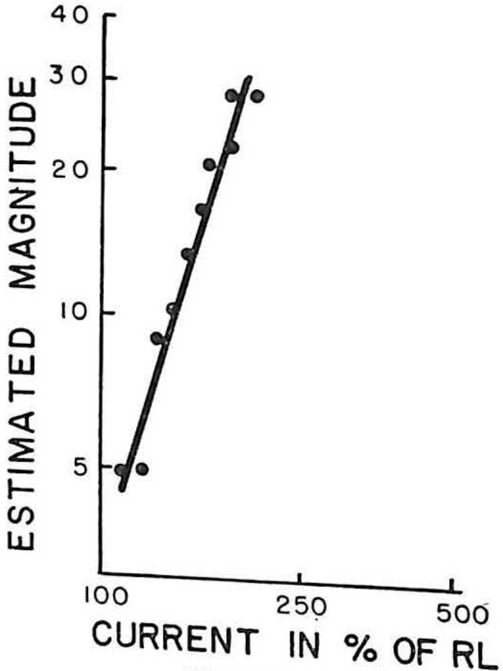

FIGURE 7
PERCEIVED INTENSITY ESTIMATES, EXPERIMENT 5

G. CONCLUSIONS

The method of magnitude estimation was found to yield data capable of being fitted by straight lines (in log-log plot). The slopes of such lines, however, are affected by such variables as instructions to the *O*s, area, and duration of stimulation. No significant difference in slope occurred as a result of reduced density of innervation or from warming the skin, although von Békésy (2) has found a faster rate of increase in subjective intensity of mechanical cutaneous vibration under similar conditions. It may be concluded that it is not possible to determine *the* (one) slope exponent for electrical cutaneous magnitude estimates, inasmuch as the exponent is significantly affected by, e. g., duration of stimulation.

H. SUMMARY

The sensitivity of data obtained by the method of magnitude estimation to variations in experimental technique was evaluated by determining changes in slope of lines fitted to median estimates in log-log plots. Instructions to the *O*s, and duration of stimulation, produced significantly different slopes in these experiments, as predicted herein. Reduction of area of stimulation tended to reduce the slope of such a line fitted to median intensity estimates, but did not affect estimates of apparent area.

Density of innervation and warming the skin did not significantly affect intensity estimates in this study, in the latter case probably due to the mode of action of the receptors. It may be concluded that it is possible to make explicit predictions of the circumstances under which the slope exponent of a line fitted to electrical cutaneous magnitude estimates should or should not change.

REFERENCES

1. BÉKÉSY, G. VON. Funneling in the nervous system and its role in loudness and sensation intensity on the skin. *J. Acoust. Soc. Amer.*, 1958, **30**, 399-412.
2. ———. Similarities between hearing and skin sensations. *Psychol. Rev.*, 1959, **66**, 1-22.
3. GARNER, W. R. Advantages of the discriminability criterion for a loudness scale. *J. Acoust. Soc. Amer.*, 1958, **30**, 1005-1012.
4. HAWKES, G. R. Skin temperature effects on vibratory thresholds. *Va. J. Sci.*, 1956, **7**, 344.
5. ———. Cutaneous communication: absolute identification of electrical intensity level. *J. of Psychol.*, 1960, **49**, 203-212.
6. ———. (*Ed.*) Symposium on cutaneous sensitivity. *USA Med. Res. Lab.* Rep. No. *424*, Ft. Knox, Ky., 1960.
7. ———. Cutaneous discrimination of electrical intensity. *Amer. J. Psychol.*, 1960. In press.
8. STEVENS, S. S. The direct estimation of sensory magnitudes: loudness. *Amer. J. Psychol.*, 1956, **69**, 1-25.
9. ———. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153-181.
10. STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.*, 1957, **54**, 377-411.
11. STEVENS, S. S., CARTON, A. S., & SHICKMAN, G. M. A scale of apparent intensity of electric shock. *J. Exp. Psychol.*, 1958, **56**, 328-334.
12. WEITZ, J. Vibratory sensitivity as a function of skin temperature. *J. Exp. Psychol.*, 1941, **28**, 21-36.

Psychology Division
US Army Medical Research Laboratory
Fort Knox, Kentucky

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 315-322.

PHENOMENAL DISTANCE IN SCENES WITH INDEPENDENT MANIPULATION OF MAJOR AND MINOR ITEMS*

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND DORIS C. DEHARDT

A. PROBLEM

A recent series of studies has dealt with determinants of perceived distance in pictures. Two of these were size of print (i. e., size of the scene as a visual target) and metric distance of the target from the eye. The *larger* the print, the nearer elements in it look to be. The *closer* the print to the eye, the nearer the scene and its elements seem to be. While these principles are probably quite obvious, Bartley and Adair (3) reduced the relations to rule. The manipulation of metric size of print and its distance from the eye while making comparisons provided information upon which to base a further procedure in dealing with target variables. Some of the same kind of information was obtained in a study by Thompson and Bartley (6).

A curious outcome of comparing photographic prints and their mirror images (prints made from reversing their negatives in the printing process) was that items appearing on the left side of the scene were perceived to be nearer than when on the right. The procedure for determining this involved comparing large prints with a small standard one, the other manipulated factor of comparison being the lateral position of the crucial item in the print. In this procedure there were four combinations possible, l-R, l-L, r-R, and r-L, in which the lower case letters stand for the small prints, and the capitals stand for the large prints. The L's (l's) and R's (r's) stand for the left-hand and right-hand positions of the crucial target items, respectively.

It has been suspected that a blanket statement that items on the left appear nearer than the same items when on the right at the same representative metric distance is too general. Not all the necessary collateral variables have been manipulated to justify it. Therefore, the present study was devised to disclose the effects of the asymmetry of the background in determining perceived distance of the crucial item. The term, crucial, for our purposes simply means the item chosen.

---

* Received in the Editorial Office on August 2, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

To make matters clearer, the following is a general description of scenes as targets: (*a*) A scene can be totally homogeneous and wholly symmetrical, as, for example, the simple picture of a prairie without any visible items aside from the terrain itself. (*b*) A scene can be symmetrical but *not* homogeneous, as the picture of the same prairie with a building on the extreme left, and a similar building on the extreme right. (*c*) Scenes can also be asymmetrical, the asymmetry and the resultant perceptual concomitants being the objects of study. A scene may be asymmetrical in its broader features, such as one in which on the left there is nothing obstructing the view of the distant horizon while on the right is a nearby clump of trees or a forest. Naturally one can reverse or otherwise manipulate this target asymmetry in any of several ways: (*a*) by use of the mirror-image procedure already described, (*b*) by use of two separate scenes which are laterally (horizontally) complementary to each other, (*c*) by shifting the camera laterally at right angles to the line of regard so as to vary the proportion of the total scene occupied by the horizon and the "forest" (*d*) by simply keeping the same camera position and revolving the camera to manipulate scene-element proportions, or (*e*) by appropriate print cropping.

The scene may not only contain major elements of spatial asymmetry such as just described, but it may also contain smaller items such as people or other small objects. These, when of concern, generally appear in the mid or foreground and occupy less of the total target area than the major elements described. Let us call such scenes two-element prints, with the major element called M, and the smaller element called S. Elements M and S may be shifted from side to side either together or independently. In this procedure, either S or M may occupy the center rather than one side of the scene. Such scenes are still asymmetrical when both are not in the center.

Actually, element M is duplex. It consists in the open portion and the filled (tree, etc.) portion.

Bartley and DeHardt (4) made a study of a further factor, namely the phenomenal distance of an item seen in the background as compared to one seen in the foreground. They found that an item seen in the distance (background) did not vary in phenomenal distance in the same manner as did objects in the midground or foreground. Part of this difference was probably due to the differences in the relative proportions of the total scene occupied in the two cases.

The purpose of the present study was to manipulate M and S independently in accord with at least some of the possibilities. Since the variable of metric size pertains not only to print elements (M and S) but also to the overall prints themselves, it has to be kept clear which is being involved. Since, according to the psychophysical procedures we use, print size as well as relative sizes of elements in the prints is involved, a distinction between the two must be made. Therefore, in using letter designations for prints, the lower case and capital letters in the first part of the statement of the factors in a given combination will refer to the small and large print sizes, respectively, and a second lower case letter (in parentheses) will refer to the position of the small item (block) in the large print.

B. Apparatus and Procedure

Four prints of the same artificial[1] outdoor scene were used. These four prints were alike in all respects except for the position and/or size of a single black wooden block which appeared in them. In three prints, a small square block appeared in the left, the center, and the right of the scene, respectively. The print containing this block in the center position is presented in Figure 1. The fourth print contained a rectangular block placed hori-

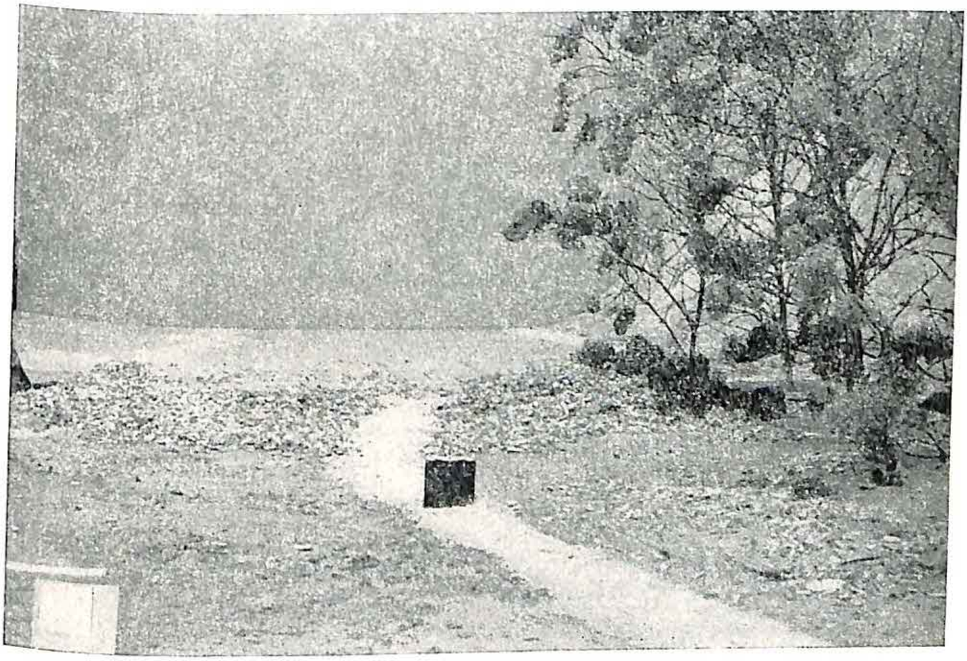

FIGURE 1
A Sample Print Used in This Study

---

[1] The scene used was a "table-top" construction styled to provide the necessary components, and when photographed, not necessarily detectable from an actual scene.

zontally in the center of the scene. The rectangle was of the same height as the square block but was five times greater in length. The use of a block dimension as a variable was only a minor feature of the present study and should be kept separate in the reader's understanding of the work reported.

For all trials, a 4 × 5 in. print was placed a fixed metric distance of 29. in. from the eye. The 8 × 10 in. prints were placed on a carriage which was mounted on a calibrated track. Before each observation, *E* positioned an 8 × 10 in. print at a randomly chosen distance along the track [the apparatus and the nature of the room used for experimentation are described at length elsewhere (3)].

*O*s were instructed that their task, on every trial, was to adjust the metric distance of the large print so that the black block in that scene appeared to be equidistant to the black block in the small print. Observations were monocular.

Ten college students, 7 males and 3 females, served as *O*s. All met a visual acuity criterion of 20/30.

The study consisted in two parts, I and II. In Part I, there were two variables, crucial item (the square block) and background, as earlier described for asymmetrical scenes in general. As was stated earlier, we shall refer to the position of the trees in dealing with the background. The position of the block was varied from left, to center, to right. The position of the trees was varied from right to left, by the mirror-image technique. Thus, there were the combinations stated below for the large and small prints, in which small case letters represent tree positions in the small print and capitals, the tree positions in the large prints, the concomitant positions of block in the large prints appearing in parentheses. The block position in the small print was always the center. All this is shown in Figure 2. Accordingly, the combinations were: r R(l), r R(c), r R(r), l R(l), l R (c), l R(r), l L(l), l L(c), l L(r), r L(l), r L(c), and r L(r).

In Part II the intent was simply to compare the small square block and the large rectangular block in the center position to determine which, if either would look nearer under that condition. Hence, only the following combinations were necessary: r R(c), l R(c), l L(c), and r L(c) for the small block and the same combinations for the large block.

Four readings were taken for each condition in both Part I and II so that all *O*s made a total of 64 trials. The conditions were presented in prearranged random orders to *O*s. All *O*s made the 64 observations in one session.

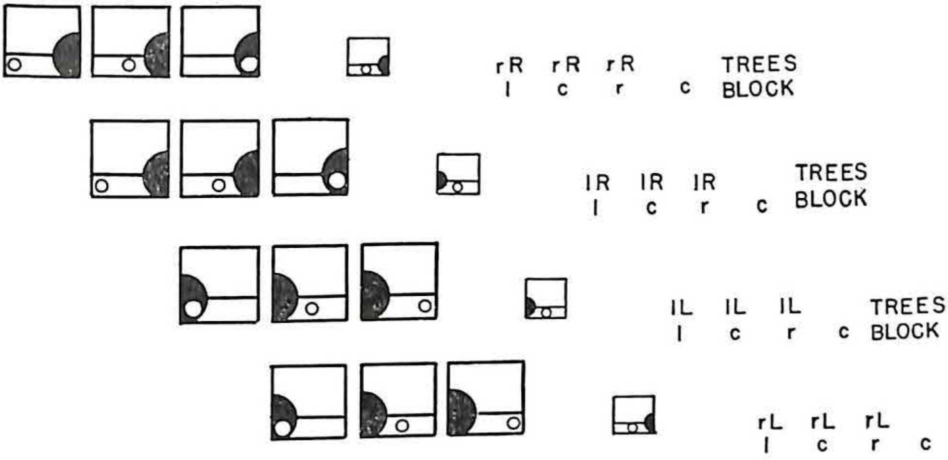

FIGURE 2
A SCHEMA OF THE TARGET COMBINATIONS USED (SEE TEXT FOR EXPLANATION)

C. RESULTS

In view of the results of previous studies (1, 4, 5), it was expected that the lateral position of the prominent objects (background objects already mentioned) in the scenes would play a large part in the determination of the apparent distance of the objects to be matched. In the present study, the trees constituted the former, the wooden blocks the latter. A problem requiring analysis was, then, that of determining the relation of the lateral position of the trees to the lateral block position in the scenes. As already pointed out, there were two locations of the trees in the scenes. The original prints contained the trees in the right portion of the scene; the mirror images, then, were scenes with trees appearing on the left. As mentioned previously, there were three block positions: l, c, r. (This was the case only in the large prints. In both small prints, the block appeared in the center of the scene. Therefore, the block's location in the 4 × 5 in. print was the same in all conditions.)

Arithmetic means were computed for each *O* for the 12 conditions in which square blocks appeared in the large prints. To determine the effects of the various tree-position combinations in the small and large prints on apparent distance of blocks in three positions, four Friedman two-way analyses of variance were computed, one each on r R, l R, l L, and r L, each test comparing three means, the *O*s' mean metric distances of the large prints when the block was at l, c, or r.

The results of the analyses were: χ_r^2 for r R = 13.4, p <.01, χ_r^2 for l R = 10.4, p <.01, χ_r^2 for l L = 3.8 (not significant), and χ_r^2 for r L

= 1.8 (not significant). The mean differences for both *R* conditions were significant, then, while neither *L* condition indicated significant mean differences.

The overall observer means for the aforementioned 12 conditions were computed and are presented in Figure 3. Note that separate curves are given for the various tree-position comparisons.

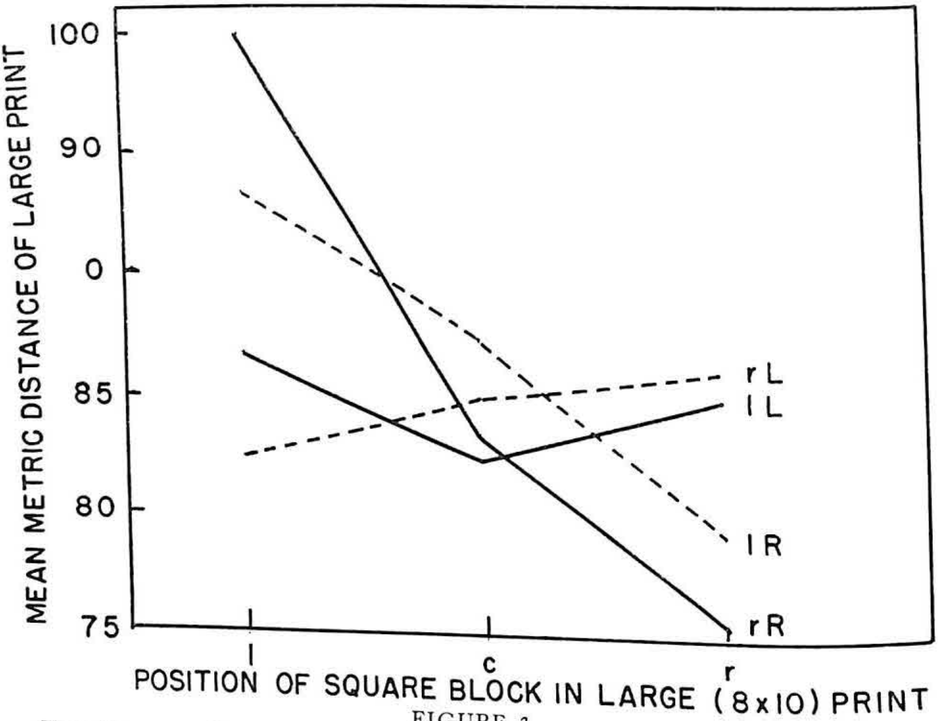

FIGURE 3

The Relation Between the Lateral Position of the Square Block and the Mean Metric Distance of the Large Print (*L* and *l* Indicate Position of Prominent Background Objects (Trees) to Left of Center in Large and Small Prints, Respectively, While *R* and *r* Indicate Trees Were to Right of Center in These Prints.)

Both statistical and graphical results indicate that when the trees were to the *right* of center in the large scene, the block appeared greatly different in distance when placed in l, c, and r positions. When, however, the trees were to the *left* of center, the apparent distance of the block varied only a little, as indicated by the very similar mean metric distances of the large prints (cf. Figure 3). And it can be seen from this figure that the position of the trees in the *large* prints is of major importance, irrespective of tree position in the small prints. Curves r R and l R are similar to each other and l L and r L are different from these but similar to each other.

It will be recalled that four comparisons for both small and large block scenes comprised Part II of this study. The analysis of results of Part II was made in the following manner. Means for all observers combined were found for the four small block and the four large block conditions. A matched-pairs t test was made on these four pairs of means, resulting in $t = 3.51$, $p < .05$. The overall mean of the large block conditions was 88.0, and small block overall mean 81.0. These results indicate that when conditions were identical in all respects except for properties of the block appearing in the scenes, the larger rectangular block appeared significantly closer than the smaller square block. This was evidenced by the fact that the prints with the large rectangular block required greater mean metric distances of the comparison prints than those with the small square blocks.

D. Discussion and Conclusions

It would appear that the general conclusion in the earlier studies (1, 4, 5), that when a target element in a three-dimensional space field is on the left, it is seen as nearer than when on the right must be qualified. It can be seen from Figure 3 that the results for trees on the right in large print conditions verify this statement, since, when a block looks nearer it must be positioned farther down the track so as to appear equidistant to the small print's block (therefore resulting in larger metric distances of the large print) than when it looks farther away. But this generalization does not hold when the trees are shifted to the left of the scene in the large print and the blocks are again varied from l to c to r. Indeed, a qualified contradiction to the conclusions of the previous studies is suggested. The r L curve is seen to *increase* slightly from l to r, the l L curve decreases, then increases (though it will be remembered from the statistical analyses that these mean differences are not significant). In other words, the left-positioned block appeared just as far away, even a little farther than the right-positioned block.

It seems it is not alone the lateral position of the target object, but its position in connection with the lateral position of the prominent background objects that determines apparent distance of the target object. The crucial object in the present study was chosen to be the small block. In the initial study in this series, the scenes as totals were dealt with. No specific item was made explicit for centering the observer's attention. The scenes were of buildings, etc., which were pretty much in the midground or even foreground and they were asymmetrically placed in the print. This asymmetry was manipulated by the mirror-image technique, and the observers were asked to match small and large prints as overall scenes during this manipulation.

The results were, of course, to the effect that items appearing on the left are nearer than when appearing on the right. These items occupied large portions of the total scene and compared in relative size to the background items (trees) in our present prints. Now, since we have shifted to a crucial item occupying quite small portions of the visual field (scene), the rule applies, only with qualifications.

It might be generalized that there is a broad left-right (horizontal) asymmetry in the phenomenal distance of components of a three-dimensional scene. Large components (forests, hills, etc.) or nearby objects such as buildings, etc., provide this left-right asymmetry, the item when on the left appearing nearer than when on the right. As smaller items, those occupying less of the total scene, are considered, they, too, tend to be seen in the same way but with qualifications. The major items' left-right position in the mid-or background enters into the determination of the phenomenal distance of the small item in accord with its left-right position. Actually, this asymmetry in perceiving the small item may never emerge if the large item is on the lift. Indeed, a slight reversal may even be suggested. In scenes which do not contain the *two* elements, but only a small and/or distant one, the asymmetry is either minimal or absent. A specific study to analyse a possible functional distinction between elements seen as *distant* in contrast to those which merely occupy small portions of the total scene has not yet been attempted.

The fact that the same components in the small and in the large prints do not function alike, or shall we say, do not have the same weight, needs further study so as to be certain that error is not being introduced through our methodology.

REFERENCES

1. ADAIR, H., & BARTLEY, S. H. Nearness as a function of lateral orientation in pictures. *Percept. & Mot. Skills*, 1958, **8**, 135-141.
2. BARTLEY, S. H. Some comparisons between print size, object position, and object position in producing phenomenal size. *J. of Psychol.*, 1959, **48**, 347-351.
3. BARTLEY, S. H., & ADAIR, H. J. Comparisons of phenomenal distance in photographs of various sizes. *J. of Psychol.*, 1959, **47**, 289-295.
4. BARTLEY, S. H., & DEHARDT, D. C. A further factor in determining nearness as a function of lateral orientation in pictures. *J. of Psychol.*, 1960, **50**, 53-57.
5. BARTLEY, S. H., & THOMPSON, R. A further study of horizontal asymmetry in the perception of pictures. *Percept. & Mot. Skills*, 1959, **9**, 135-138.
6. THOMPSON, R. W., & BARTLEY, S. H. Apparent distance of material in pictures associated with higher order meanings. *J. of Psychol.*, 1959, **48**, 353-358.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 323-332.

CERTAIN CHROMATIC AND BRIGHTNESS CHANGES ASSOCIATED WITH RATE OF INTERMITTENCY OF PHOTO STIMULATION*

Department of Psychology, Michigan State University

S. Howard Bartley and Thomas M. Nelson[1]

A. Problem

Intermittent photic stimulation produces a host of visual effects. Despite this, only a very few have been extensively studied, the chief one being the conditions for just eliminating brightness fluctuations (flicker). This point has been called the fusion point and the rate of stimulus intermittency at which this point is reached is called critical flicker frequency (CFF). At various times, effects produced by subfusional pulse frequencies have been incidentally mentioned. Bartley (1) varied subfusional pulse rates and noted the phenomena produced. He reported not only various kinds of flicker in the test area, but flicker in the surrounding field, gamma movement, and ocular tension, etc. One of the results stemming from this brief survey was the study of brightness enchancement, the increased brightness produced by low rates of intermittent stimulation over the brightness produced by a steady presentation of the same intensity (1, 2, 3).

One of the characteristics of the field produced by intermittent stimulation is a greatly changed surface character (3). This imposes some difficulty in obtaining a brightness match between the field produced by intermittent stimulation and a standard field alongside it. One feature of the field produced by intermittent stimulation is glitter. A peculiar form of surface quality is produced in a compound target in which the intermittent portion lies within the steady portion as is the case in using a ring-annulus target. If one begins with a low rate of intermittency which produces the experience of alternations of light and dark, one can gradually increase the rate of intermittency and produce a rather unique phenomenon. As rate is increased the alternations of light and dark do not follow each other in a regular fashion but tend to become temporally jumbled. This shift to irregularity changes still further as rate is further increased until both

---

* Received in the Editorial Office on August 15, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.

phases actually appear simultaneously. This sounds absurd, but it is accomplished by the light phase lying behind the dark phase and showing through it as if the darkness were a textured translucency. This amounts to a *temporal alternation* turning into a *spatial simultaneity,* an example of what would seem to be a very fundamental feature of the relations between impingement and sensory response (3).

Another salient phenomenological feature of intermittent stimulation lies in the domain of color. Either whole spectrum stimuli produce colors, when only spectral stimuli would be expected to do so, or when spectral stimuli are used, the hue and saturation may change in accord with intermittency rate of the stimulus.

Fechner, with a certain arrangement of lowly reflective and highly reflective sectors ("black" and "white") on a rotating disk, obtained a pattern of blue and yellow. Parsons (9) mentions Helmholz's observation that, under some conditions, the leading edge of the highly reflective ("white") sectors of a rotating disk appeared reddish, and the following edges bluish.

In the investigation of brightness enhancement, the chromatic and glitter effects have been disturbing features. In making brightness comparisons between two fields, naturally, the two should have the same chromatic appearance whether this is a "good" white, or some off-white, or even a fully saturated color. Since rate of stimulus intermittency is a factor which so definitely varies chromatic character of the response, it can hardly be expected that experimenters will be able to make adjustments so that hue will always be matched in the two targets. This must be taken into account in evaluating studies of brightness enhancement. The present study was meant to point out this dependency and to provide at least some qualitative information on it.

B. Apparatus and Procedure

Stimulus intermittency was produced by an episcotister, using a two-open-sector disk producing a PCF (pulse to cycle fraction) of ⅛. The luminous source was a 50 cp automotive head-light lamp fed from storage batteries, the voltage regulated to provide a target luminance of 1400 c/ft^2. For the work we used the full luminosity and three lower levels of intensity, given by three filters, reducing intensity by logarithmic steps. The target size was 9/16 x 11/16 inches and was viewed monocularly from a distance of 30 inches. Wratten filters were used. They were No. 70, "dark red," with a cut off at 650 mμ, (dominant wave length 678 mμ); No. 71A, "red," with a cut off at 610 mμ (dominant wave length range of 630-

640 mμ); No. 72, "orange," with a cut off at 590 mμ (dominant wave length of 610 mμ); No. 73, "dark yellow green," with a range of 560 to 576 mμ (dominant wave length 576 mμ); and No. 76, "dark violet," with a range from 400 to 470 mμ (dominant wave length 448 mμ).

These filters, of course do not transmit the same overall amount of energy. The following amounts of luminous transmissions are: filter 70, 0.68 per cent; 71A, 0.90 per cent; 72, 1.3 per cent; 73, 1.41 per cent, and filter 76, 0.015 per cent. For our purposes a strict equal energy comparison was hardly necessary. The effect that energy involment has upon sensory end-result was sampled by using the Wratten neutral filters already mentioned in conjunction with the color filters, so as to deal with four logarithmic steps from about 1400 c/ft^2 to 1.4 c/ft^2 minus the color filter luminous transmittance.

The observer's task was to note changes in brightness and color as he manipulated the intermittency rate of the stimulus (revolution of the episcotister disk) at will. He began with the disk stationary and then varied the intermittency rate upwards to the desired limits. The observer was to indicate the direction of change in brightness, where brightness appeared to be maximum if any marked maximum showed up, where brightness seemed to stabilize, or to return to the original level. Furthermore, the observer was to indicate critical flicker frequency and to go beyond this rate to see whether any noticeable change occurred above the CFF. This range in Tables 1 and 2 is called "above CFF."

At least two sets of readings were made on the points just listed for both ascending and descending trials for each of the five color filters at each of the four intensity levels, and for CFF.

Another complete set of trials was obtained with the observer noting color changes in about the same manner as for obtaining the brightness changes. The two authors served as the principal observers.

Obviously no trials involved fixed standards of comparison. All were made by noting *changes* of *more* or *less* as stimulus intermittency was varied. There was no intention of developing a precisely quantitative statement of the sensory outcome but rather of making a broad initial survey permitting certain inferences in regard to the factors at work.

C. RESULTS

The results are shown in Tables 1, 2, and 3. It will be seen that the first changes in both brightness and in color occur either as soon as the target is made intermittent or very shortly afterwards, for all filters and for all intensities used. In this, both of the observers were essentially alike. In all but three intensity-spectral conditions for Observer *B* and four for *N*, the change was a drop in brightness. In most cases this drop was reversed and some enhancement took place as photic pulse rate was raised. Sometimes this increase in brightness was slight, sometimes marked, depending upon the part of the spectrum being used. In many cases, no brightening maximum could be detected by the qualitative methods we used. In cases in which brightening occurred it tapered off at some rate below CFF and at other times a second increase in brightness appeared at rates above CFF. It must be remembered that no standard target for comparison was used and therefore the extent of brightening could not be determined. Actually when an initial drop in brightness at low intermittency rates of stimulation occurred, the brightening may never have exceeded the level of the steady target. The three filters producing this supra-CFF increase in brightness for Observer *B* were 71A, 72, and 73, the three in the intermediate portions of the spectrum. In the case of Filter 73, the supra-CFF increase occurred at Intensity 3 as well as at Intensity 4, the level at which the other two filters produced it. Observer *N* reported supra-CFF brightening only at Intensity 4 for Filter 73.

As was said, color changes began to take place almost immediately when slow intermittency was introduced. As rate was increased, the color change, a kind of fading of hue (desaturation) or a shift in hue became more marked. To note their nature, see Table 3. In a number of instances a *color* change occurred above CFF. This change and the change in *brightness* above CFF are two of the more unique observations we have made, since, to our knowledge, nowhere in the literature has a supra-fusional change been established, or perhaps even reported. Possibly if such changes were ever observed they were later supposed to be artifacts. We can imagine no artificial factors in our conditions that would account for the modest but indubitable changes we obtained. The color changes were more frequent and easier to observe than the brightness changes, and occurred under more different conditions than brightness changes. Under most conditions, a detectable maximum for the color changes below fusion was observed. No color filter failed to produce it for some intensity or frequency condition.

D. Discussion

While this study had for its prime purpose the consideration of the role of intermittency rate in changing both brightness and chromatic qualities, the findings although not precisely quantitative do suggest that a comparison

TABLE 1
RELATIONS BETWEEN BRIGHTNESS, COLOR, INTENSITY AND STIMULUS RATE
(Observer *B*)

		Brightness changes				Color changes				
Filter	Intensity	First change	Brightening range	Terminal below CFF	Above CFF	First	Maxi-mum	Terminal below CFF	Above CFF	CFF
76	4	B 2.0	2.0– 4.8	17.8		5.0	8.0	18.5	39.0	31.5
	3	D 5.0	1.5– 3.5	9.8		0	5.0	10.6		15.6
	2	D 3.0		no changes		0	4.0	8.5		15.6
	1	D		constant decrease		no C	no C	no color difference		14.6
73	4	D 0.5	15.5–27.5	27.5	B 53.0	2.2	10.0	19.0	43.0	37.0
	3	B 1.9	1.9–37.0	37.0	D 97.0	3.1	9.5	14.5		29.4
	2	D 3.5				3.2	9.5			23.1
	1	D 0+				no C				17.0
72	4	B 3.6	3.6–13.0	18.8	44.0	1.9	5.5	17.0	44.2	34.4
	3	D 3.5		7.0		2.1	11.0	11.0		27.3
	2	D 3.9				3.8	10.0	10.0		19.8
	1	D 0+				2.8	2.8	2.8		14.6
71A	4	D 2.8	8.8–16.0		39.5	2.8	6.2	24.0	40.0	34.5
	3	D 2.1				2.0	18.5		39.5	27.0
	2	D 5.0				3.8	9.2	9.2		19.5
	1	D 0+				3.0	3.0	3.0		13.6
70	4	D 2.9				1.0	7.0	21.0		31.4
	3	D 2.1				3.5	9.9		30.7	24.8
	2	D 4.0				3.0	8.5	8.5		16.5
	1	D 0+				3.0	3.0	3.0		13.8

TABLE 2
RELATIONS BETWEEN BRIGHTNESS, COLOR, INTENSITY AND STIMULUS RATE
(Observer *N*)

		Brightness changes				Color changes				
Filter	Intensity	First change	Brightening range	Terminal below CFF	Above CFF	First	Maximum	Terminal below CFF	Above CFF	CFF
76	4	B 2.0	1.5– 3.5	25.0		2.5	12.7	30.5		37.9
	3	D 2.5	1.5– 3.5	15.0		1.0	2.5	15.0		27.2
	2	no E				0	none	7.7		16.2
	1	no E				2.5		12.0		12.5
73	4	D 2.0	max. 6.5	17.7	B 56.5	4.0	13.2	30.0		44.3
	3	B 2.2	2.2– 6.8	max. 4.4		2.5	6.8	18.4		37.1
	2	D 2.0	2.0– 3.0			0		13.0		25.6
	1	D 0								17.0
72	4	B 7.0	max. 11.0	14.5–17.0		3.5		11.0–14.0	40.5	39.4
	3	D 3.6	3.6–13.0			1.9	5.5		42.2	34.4
	2	B 2.0				0		10.0		22.2
	1	D 0				1.5	1.5–48	4.8		15.4
71A	4	D 2.0		16.5–20.0		1.5	8.7		44.5	40.5
	3	D 2.8	8.8–16.0	max. 10.0		2.8	6.2		28.0–40.9	34.5
	2	B 2.4				0	11.4			22.7
	1	D	5.2			2.5	none	10.0		14.1
70	4	D 2.0	11.0–14.0			1.0	5.75–8.75		36.5	37.8
	3	D 2.9				1.0	7.0	21.0		31.4
	2	D 4.0	8.0–16.0			0		11.2		18.6
	1	D 2.5	no E			2.8		9.8		11.6

TABLE 3
DESCRIPTION OF COLOR CHANGES AS RELATED TO STIMULUS FREQUENCY AND INTENSITY
(Observer *B*)

Filter	Intensity	Color for steady field	Fading and/or other changes
76	4	bluish violet	To yellowish green, to glittering whitish purple or violet, returning yellow with greenish tinge to blue violet.
	1		Only a graying of original color.
73	4	greenish yellow	To golden, no glare or glitter, changes to amber at 31.5, becomes stronger till terminal fading. Post-fusional color is about same as original, a zinc yellow.
	1		No fading, yellowish white loses brightness rather than saturation.
72	4	clear orange	Yellow to whitish glaring yellow, at maximum yellow returns and becomes stronger, becoming an orange. Above fusion the orange brightens, almost appearing as desaturation.
	1		Too dim to judge color changes.
71A	4	red	Yellowish red turning to orange-amber becoming amber with reddish tinge. Above fusion red goes toward ochre or amber again.
70	4	brilliant red	To amber to orange, red gradually returns, yellow tint leaves. Post-fusion is red.
	1		Stays red after losing original saturation.

between them and neurophysiological findings might be possible. Various workers have used spectral stimulation in the study of retinal responses of single units or of the overall retina (ERG), in the study of optic nerve discharge, in the response of the lateral geniculate body and of the optic cortex itself. Definitive results are beginning to accumulate in this area.

Chang (4) has reported faster central conduction rates for the long wave lengths of the spectrum than for the short. Madsen and Lennox (8) report various differences in latency, shape, and amplitude of the cortical responses to different wave lengths. Cortical *on* responses were most pronounced to the "yellow" and "red." The amplitude was greater to "blue" than to "green" with stimuli adjusted in intensity to produce equal sized retinal responses. The latency to "blue" was longer than to "green," "yellow," or "red" when intensities were adjusted to produce equal retinal latencies. Lennox (5) found that spectral stimuli at intensities producing equal response latencies at the lateral geniculate resulted in different latencies at the cortex. The latencies to long-wave-length stimuli were shortest, longer to mid wave lengths and longest to the short wave lengths. The findings were interpreted as suggesting that impulse transmission between geniculate and cortex differs for different wave lengths.

Lennox (6) found in single fibers of the optic tract that pure *on* fibers were more likely to be short-wave-length sensitive, and tended to belong to the slower conduction velocity groups. In the *on-off* samples, only a very few long-wave-length sensitive fibers were found; and they belonged in the fast conduction groups. *On* fibers were in general longer in latency, productive of fewer spikes, and a lower discharge frequency than other types of fibers. The *on-off* were, of course, just the opposite in characteristics. Spontaneous activity was present in the *on-off* fibers and almost entirely absent in the *on* fibers. Slowly conducting fibers responded more actively to the shortest wave lengths than to the longest, by a factor of 2.5 times. "Fast" fibers responded more actively to the longest wave lengths than to the shortest. These two groups did not overlap in their values.

Lennox (7) also studied *off* responses in the cat. Such responses were of three kinds: those that were intensity-sensitive, those not sensitive in this way, and those that manifested mixed sensitivity. When pulses of previously constant intensity were shifted up or down there was a nearly linear relationship between number of spikes and flash duration in the 1 to 40 msec. range. Latency tended to remain constant except at the extremes of the range of intensity or duration. A number of other characteristics were also reported but the foregoing will serve to illustrate the kinds of variables in both stimulation and response that are found in studying the optic pathway.

What is the relation between these findings and the sensory findings we have just described? Naturally we have no basis for particularizing to any great detail, but it would seem that certain suggestions are in order.

If, as it has been reported, various functional discharge units of the retina, geniculate, and cortex behave differently to diverse parts of the spectrum, these differences in response may be affected by repetition rates of intermittent stimulation. The neurophysiological results that have been reported include differences in latency and discharge rates in the train of impulses constituting responses to short photic pulses. Since there are only two differentiating factors that can operate in retinal discharge, namely *which* units are involved, and the *manner* in which they discharge, the sensory end-result of color must originally stem from these two factors. Historically attention has been focused on the first of the two factors. This issue has been discussed in terms of how many different color units there are in the retina, implying that these, no matter how they behaved, would send messages up certain channels and not others to the visual cortex and the cortex would respond, at least in part, in accord with where the messages came from. Now we are aware of the specific impulse discharge characteristics involved when spectral stimuli are used. Hence we are now in a position to consider how *temporal* characteristics of discharge may convey the necessary information to the cortex, without asking how the discharges are produced. This mode of inquiry is closer to other forms of study of both the stimulus, and the sensory end-result. It is this form of inquiry that would tend to integrate the study of color and the study of flicker and other more obviously temporally controlled sensory end-results.

Heretofore the temporal characteristics of retinal discharges and cortical inputs, or cortical responses themselves have been most closely examined for relating them to CFF's, but since it has been found the different parts of the spectrum set up neural responses consistently differing in temporal characteristics, anything that alters overall temporal patterning of these spectrally differentiated responses, such as repetitive stimulation, ought to affect chromatic sensory responses in some way.

If some portions of the spectrum produce neurophysiological responses with long latencies, and other portions produce responses with short latencies, these responses should be differentially affected if repeated stimuli are used. The responses with long latencies would be affected sooner and to a greater extent as intermittency rate is increased from minimal values. Lennox found that at the cortex, latencies of responses to long wave lengths were shortest. This means that the seeing of red should be inffuenced less by intermittency rate being increased, than would the seeing of blue.

How do these facts relate to the data in Tables 1, 2, and 3? These tables show by their absence of readings on the left hand side, that the brightness of red was affected less than the brightness of the other hues. Table 3 indicates that the two long-wave-transmitting filters produced sensory end-results which changed chromatically far less than the end-results from use of the other filters. Even when the red was observed to change it changed less radically than the other colors. In case of filter No. 70, the color remained red throughout, and this filter was the one transmitting only the very longest wave lengths.

Neurophysiological studies have dealt with not only the factors just mentioned, but have disclosed some differences between the roles of the *on, on-off,* and the *off* units.

It was found that the slowly conducting units responded most actively to the short wave lengths. This is to say then, that the blue-producing stimuli activate the more slowly conducting units. This too, would suggest that repetition rates of stimuli would affect the blue-producing stimuli most. This again tallies with our findings shown in the tables.

Notice should be taken of what may appear to be an exception. It is the set of findings for the lowest level of intensity for the "violet" filter (76). There it is stated that no color differences were found. The findings here were that there was very little color to begin with. There was so little that the only change was desaturation or dimming into no color at all. Hence, when this level is considered in light of this, no contradiction exists.

E. Summary

The present paper is a study of the relation between color and brightness of spectral targets and the intermittency rate of stimulation. These findings are discussed in light of expectations derived from recent neurophysiological work.

References

1. Bartley, S. H. A central mechanism in brightness discrimination. *Proc. Soc. Biol. Med.*, 1938, **38**, 535-536.
2. ———. Some factors in brightness discrimination. *Psychol. Rev.*, 1939, **46**, 337-338.
3. ———. Some effects of intermittent photic stimulation. *J. Exper. Psychol.*, 1939, **25**, 462-480.
4. Chang, H. T. Functional organization of visual pathways. *Res. Pub. Ass. Nerv. & Ment. Dis.*, 1950, **30**, 430-453.
5. Lennox, M. A. Geniculate and cortical responses to colored light in the cat. *J. Neurophysiol.*, 1956, **19**, 271-279.
6. ———. The on responses to colored flash in single optic tract fibers: correlation with conducting velocity. *J. Neurophysiol.*, 1958, **21**, 70-84.
7. ———. Single unit off responses to brief flashes in cat's optic tract. *J. Neurophysiol.*, 1959, **22**, 88-97.
8. Madsen, A., & Lennox, M. A. Cortical and retinal responses to colored light flash in anesthetized cat. *J. Neurophysiol.*, 1955, **18**, 412-424.
9. Parsons, J. H. An introduction to the Study of Color Vision. Cambridge: Cambridge Univ. Press, 1924.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 333-337.

SOME RELATIONSHIPS AMONG PERSONAL ADJUSTMENT SELF-RATINGS, SELF-SCORES, AND ASSIGNED "AVERAGE" SCORES*

Department of Psychology, University of California, Los Angeles

---

GEORGE F. J. LEHNER

A. INTRODUCTION

In a previous study (2) the author investigated the relationship between personal adjustment scores and projected "average" scores. The subjects in this previous investigation answered items on the California Test of Personality (*a*) for themselves and then (*b*) as they believed the "average" person would answer them. It was found that regardless of the individual's own score or level of adjustment, the projected "average" score was lower than his own score. In short, the subjects viewed the "average" person as less well adjusted than themselves.

The purpose in this investigation is to determine the relationship of an additional variable, namely the self-rating of personal adjustment and its relation to self-scores and projected "average" scores. This approach will make it possible to discover whether poorly adjusted individuals rate themselves differently from average or well-adjusted individuals; it will also illustrate the relative position of self-*ratings* with respect to self-*scores* and the projected "average" scores.

On the basis of prior research one might expect to find a relationship between the three variables here studied. Sears (4) states that self-evaluation and evaluation of others are related variables, and that the less insight an individual had to a particular trait of his own, the more often he attributed that trait to others. Chodorkoff (1), investigating the relationship of self-perception, perceptual defense, and adjustment, found that individuals who are inaccurate in their perception of the environment are also inaccurate in their self-perceptions, and that the less accurate the self-perception, the poorer the personal adjustment.

B. PROCEDURE

The California Test of Personality, Adult Series, was administered to 132 students (68 male and 64 female, average age = 21.5) in two evening

---

* Received in the Editorial Office on August 17, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

classes in elementary psychology. No time had yet been devoted in these classes to the topic of personality adjustment or personality measurement.

The test was given twice to the same group of students. In the first administration, each student answered the test items for himself. They were told that they were to answer the questions as honestly as possible and that they did not have to give their names. They did, however, mark their papers with an identifying code so they could be matched later to papers from a later administration. A second administration followed one week later. At this time, they were given the same test with instructions to answer each item as they believed the majority of persons (representing the statistical "average") would answer.

After the subjects completed the test for the second time they were given a sheet of paper with a chart marked in deciles from 0-100; "50" was listed as "average," over 50 was listed as "better than average," and less than 50 as "poorer than average." The subjects were then asked to place a check mark at the decile score which they believed represented their own "total, general, over-all level of adjustment."

C. Results

An initial analysis on the basis of sex was made for the relationship between "self" scores, self-ratings, and projected "average" scores (Table 1).

TABLE 1

	Men	Women	Total
N	68	64	132
Mean self-score	38th percentile	47th percentile	43rd percentile
Mean "average" score	18th percentile	24th percentile	22nd percentile
Mean self-rating	61st percentile	62nd percentile	62nd percentile
Critical ratio — mean self-score — mean "average" score	5.4*	6.7	6.0
Critical ratio — mean "average" score — mean self-rating	14.5	13.1	13.8
Critical ratio — mean self-score — mean self-rating	8.7	5.2	6.9

* All critical ratios are significant at the 1 per cent level of confidence.

1. *Men*

As indicated in Table 1, the mean self-score for men is at the 38th percentile. The mean projected "average" score, as in our previous study, fell below the self-score and is at the 18th percentile. The mean self-rating, on the other hand, is at the 61st percentile. The differences between all three of these means are significant at the 1 per cent level of confidence.

2. *Women*

Results for women are approximately the same as those for men. The mean self-score is at the 47th percentile, while the projected "average" mean score is again low—at the 24th percentile. The women, like the men, rated themselves as better adjusted than their scores for "self" would indicate; the mean score for self-ratings is at the 62nd percentile. Again, the differences between means are significant at the 1 per cent level of confidence.

3. *Total Group*

It is apparent that there are no real differences between men and women in the results obtained. A check was made on the largest discrepancy between men's and women's scores, i.e., the difference in mean self-scores, and it was found non-significant at the 1 per cent level of confidence. Thus the scores for both sexes were pooled and a similar analysis was made.

The total mean score for "self" is at the 43rd percentile, the mean score for "average" is at the 22nd percentile, and the mean self-rating is at the 62nd percentile. The differences between these means are significant at the 1 per cent level of confidence.

Another analysis of the data was made. "Self" scores on the California Test of Personality were divided into three groups: those indicating better than average adjustment, average adjustment, and poorer than average adjustment. Scores 70 and over represented "better than average" adjustment, scores from 40 through 69 were taken as the limits for "average" adjustment, and scores 39 and below represented "poorer than average" adjustment. These were then called the High group, the Medium group, and the Low group. The three groups were compared to determine whether they *rated* themselves differently or gave different projected "average" scores. Results show that although all three groups rated their own adjustment as relatively high while scoring the "average" adjustment as relatively low, there were significant differences in this tendency for the three levels of adjustment as based on self-scores. In a 3×3 contingency table, comparing the three levels of adjustment against the three levels of self-rating (high, medium and low adjustment), a X^2 of 15.24, significant at better than the .01 level of confidence, was obtained. A proportionately larger number of individuals in the well-adjusted (High) group rated themselves as well-adjusted; the Medium group ("average" adjustment) were divided between rating themselves as well-adjusted or average-adjusted, reflecting the general findings for the entire group; a proportionately larger number of the Low group ("poorer than average" adjustment) rated themselves as having average adjustment. In general, the higher the self-score on adjustment the higher the self-rating.

In a 3 × 3 contingency table comparing the three levels of adjustment with the three levels of projected "average" scores, a X^2 of 11.24, significant at better than the .05 level of confidence, was obtained. A proportionately larger number of well-adjusted individuals scored the "average" person as being average or well-adjusted. Individuals in the Medium and Low groups had projected "average" scores lower than the High scoring group, i.e., scores typical of the group as a whole. In general, the high-scoring (superior adjustment) persons projected higher scores for the "average" person.

D. DISCUSSION

The relationship obtained in the present investigation between "self" scores and "average" scores is almost identical to that found in a previous study of these two variables (2). Regardless of their own scores, college students perceive the "average" person as less well-adjusted than themselves, and in fact, as a poorly adjusted person.

Although all individuals at all three levels of adjustment reflect the group tendency of rating themselves as relatively well-adjusted and the "average" person as poorly adjusted, this tendency varies with the different levels of adjustment. In particular, the well-adjusted individual has a greater tendency to perceive both himself and the "average" person as being relatively well-adjusted.

It might be noted here that while *degree* of adjustment, as measured by the California Test of Personality, is a variable in this research, the *relations* of this degree of adjustment among self-ratings, self-scores, and projected scores is probably not affected by the question raised by Lindgren (3) on the acceptance of the California Test of Personality scores as measures of adjustment. He suggests that they may actually be measures of "culturally acceptable defensiveness" or "fear of nonconformist tendencies." The relationships between the way we rate ourselves, score for ourselves, or assign scores for others, might be the same regardless of the particular behavior or traits so investigated. It may be that these data point to a principle of interpersonal perception which might be called the "I-high, you-low" phenomenon, an attitudinal form of self-upmanship. If this "I-high, you-low" attitude is typical of the average person, it is interesting to speculate about its effects on interpersonal relationships.

E. Conclusions

1. College men and women have higher self-ratings than actual self-scores on a personality test, and both self-ratings and self-scores are considerably higher than their projected adjustment scores for the "average" person.

2. The higher the self-score on adjustment (when the total group was divided into High, Medium, and Low adjustment groups) the higher the self-rating.

3. The High self-scoring group (superior adjustment) projected higher scores for the "average" person.

4. It appears from this and related research done by the author that one might be dealing here with a "I-high, you-low" phenomenon of self-upmanship in interpersonal perceptions.

References

1. Chodorkoff, B. Self-perception, perceptual defense, and adjustment. *J. Abn. & Soc. Psychol.*, 1954, **49**, 508-512.
2. Lehner, G. F. J. Personal adjustment scores and assigned "average" scores. *J. of Psychol.*, 1956, **42**, 227-236.
3. Lindgren, H. C. The development of a scale of cultural idealization based on the California Test of Personality. *J. Educ. Psychol.*, 1952, **43**, 81-91.
4. Sears, R. R. Experimental studies of projection: I. Attribution of traits. *J. Soc. Psychol.*, 1936, **7**, 151-163.

Department of Psychology
University of California
Los Angeles 24, California

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 339-347.

A STUDY OF INCIDENTAL LEARNING*

New York University and Barnard College of Columbia University

M. E. Tresselt and M. S. Mayzner[1]

A. Problem

Incidental learning has been defined as "learning which apparently takes place without a specific motive or a specific formal instruction and set to learn the activity or material in question" (1, p. 210). Thus, in the incidental learning of verbal material, typically *S*s are instructed to react to or operate on certain dimensions of the stimulus material and are then tested for retention on other dimensions. For example, Myers (2) asked *S*s to count all letter *O*'s distributed among a group of other letters and then measured the recall of all other letters present. Most recently the problem of incidental learning has been subjected to an intensive analysis by Postman and his associates (3, 4, 5), employing a variety of techniques for the presentation of the stimulus material to the intentional and incidental learning groups.

The present study takes as its point of departure certain of the theoretical notions advanced by Postman and his associates to account for some of the phenomena of incidental learning. Thus, Postman has employed as an orienting hypothesis: that incidental learning is a positive function of the number of differential responses evoked by the stimulus materials.

This study attempts to manipulate the number of differential responses evoked by the stimulus by varying the operations to be performed on the stimulus by *S*. Specifically, *Ss* are presented with a list of words and instructed to perform certain operations on the words, i. e., one group is instructed to cross-out all vowels, another group to copy the list, and a third group is instructed to judge the words for inclusion in the concept "economic." Following completion of these tasks, *S*s are tested for their retention of the word-list. Interest is focused primarily on the effects on retention of the different operations performed on the words. In particular, the different operations may be ordered in terms of the degree to which *S* engages the stimulus word qua word. That is, at one extreme where *S* is instructed to cross-out vowels, the stimulus word as such, may be treated merely as a

---

* Received in the Editorial Office on August 27, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by a grant made to M. E. Tresselt by the New York University Arts and Science Research Fund.

series of letters which *S* searches for vowels, while with the judgment operation, an explicit differential response to the word itself and its associated meanings is required. Perhaps somewhere between these two extremes lies the operation of copying the word. It sems plausible to assume then that as the degree of *S*'s engagement with the stimulus words increases, there will be a corresponding increase in the number of differential responses evoked by these words, and therefore a greater degree of incidental learning should occur.

B. METHOD

1. *Subjects*

Two hundred *S*s obtained from undergraduate psychology courses at New York University were employed in the study.

2. *Materials*

The stimulus material presented to *S*s consisted of a list of 100 words (shown in Table 1). These 100 words were selected from a larger sample of 600 words employed in an earlier scaling study (6). In this earlier scaling study, 100 *S*s were asked to judge or check each of 600 words as belonging to none, 1, 2, 3, 4, 5, or 6 concepts, the concepts being theoretical, economic, aesthetic, social, political, and religious. By summing the number of checks for each word for the 100 *S*s, 600 word-profiles were obtained, which showed the frequency with which each word was judged as belonging to each concept. Using only the scale values (i. e., the number of checks for a given concept) for the economic concept, 100 words were selected for this study, so that the words appeared as nearly as possible at equal intervals along the scale of checkability. That is, the words in Table 1 are arranged in order of increasing checkability (as obtained in the earlier scaling study) and the associated scale values (shown under Column I) range from 1 to 100, in equal steps of one unit, except in a few cases, where there is a two- or three-unit jump.

3. *Procedure*

The *S*s were presented with the words shown in Table 1, completely randomized as to initial scale value, on mimeographed sheets. Of the 200 *S*s tested, 100 *S*s were instructed to judge the degree to which each word belonged or could be included in the concept "economic," using a seven-point scale. The number "one" was to be used if they judged the word as not belonging or belonging to a very small degree to the concept; the number "seven" was to be used if they judged the word as belonging completely or to a

TABLE 1
WORDS AND ASSOCIATED SCALE VALUES USED IN THE STUDY

	Scale values			Scale values	
Words	I*	II**M	Words	I*	II**M
1. Poem	1	1.17	51. Liquidate	51	4.57
2. Horizon	2	1.86	52. Partner	52	4.84
3. Priest	3	1.43	53. Volume	53	4.61
4. Kick	4	1.21	54. Union	54	5.70
5. Blend	5	2.16	55. Power	55	5.17
6. Skeptic	6	2.08	56. Report	56	3.67
7. Classical	7	1.57	57. Policy	57	5.14
8. Loyalty	8	2.34	58. Quality	58	4.61
9. Counterpoint	9	2.09	59. Personnel	59	5.20
10. King	10	2.68	60. Save	59	5.06
11. Banquet	11	2.30	61. Career	61	4.65
12. Violent	12	1.73	62. Ransom	62	2.91
13. Illumination	13	2.29	63. System	63	5.43
14. Investigate	14	3.48	64. Utilize	64	5.02
15. Civil	15	3.33	65. Reward	65	3.43
16. Appeal	16	2.80	66. Boss	65	4.99
17. Execute	17	2.85	67. Statistics	67	5.41
18. Evolution	18	3.01	68. Work	68	6.03
19. Judgment	19	4.05	69. Value	69	5.59
20. Tea	20	1.97	70. Office	71	4.37
21. Trunk	21	1.28	71. Bonds	71	6.42
22. Comparison	22	3.02	72. Trust	72	5.66
23. Chair	23	1.53	73. Department	73	3.45
24. Form	24	2.53	74. Increase	74	4.20
25. Travel	25	3.27	75. Client	75	4.54
26. Method	26	4.16	76. Demand	76	5.40
27. Selfish	27	2.94	77. Appropriation	77	5.66
28. Entry	28	3.28	78. Competition	78	6.12
29. Win	29	3.05	79. Margin	79	5.22
30. Blanket	30	1.64	80. Cartel	80	5.54
31. College	31	3.98	81. Consumption	82	5.87
32. Object	32	2.34	82. Machinery	83	4.95
33. Drive	33	3.89	83. Petroleum	83	4.66
34. Experience	34	3.62	84. Coal	84	4.07
35. Free	35	3.46	85. Rich	85	5.10
36. Society	36	5.13	86. Imports	86	6.30
37. Restriction	37	3.97	87. Management	87	5.92
38. Concrete	38	3.14	88. Rates	89	5.78
39. Socialism	39	4.56	89. Receipt	89	5.10
40. Standards	40	5.33	90. Budget	92	6.44
41. Charity	41	3.26	91. Expenditure	92	5.85
42. Bridge	42	2.90	92. Sale	92	5.80
43. Area	43	2.95	93. Monopoly	93	6.40
44. Company	44	5.37	94. Depreciation	94	5.58
45. Balance	45	5.17	95. Invest	95	6.54
46. Flood	47	3.41	96. Trade	96	6.36
47. Organize	47	5.46	97. Assets	96	6.20
48. Wear	48	2.23	98. Finance	98	6.67
49. Deal	49	4.62	99. Wages	99	6.35
50. News	50	3.60	100. Industry	100	6.32

* Column I gives the scale values (frequency of checkability) for the words from the earlier scaling study.

** Column II gives the scale values for the words obtained in the present study.

very strong degree to the concept. The intermediate numbers "two" through "six" were to be used to express intervening degrees of belonging or inclusion in the concept. The results of these judgments are given under Column II in Table 1. The scale values shown were calculated by averaging the ratings (1 through 7) given by the 100 *S*s for each word. Immediately after the judging task, 50 of the 100 *S*s were asked to recall as many of the 100 words as possible, while the remaining 50 *S*s were asked for their recall performance approximately 48 hours after the initial judging task. All *S*s were given 10 minutes for their recall task, and had no prior knowledge that they would be required to recall the words they had judged. The remaining 100 *S*s were divided into two groups of 50 *S*s each. One group was instructed to cross-out all the vowels in each word—*S*s were told *E*s were interested in developing a perceptual test. The other group was instructed to copy each word— *S*s were told *E*s were interested in obtaining samples of handwriting. Both were then immediately tested for their recall of the words. Ten minutes was again allowed for the recall test, of which *S*s had no prior knowledge. All testing was carried out during regular class periods, which allowed ample time for the initial and recall tasks. None of the *S*s when questioned expressed any suspicion that a subsequent recall test was planned.

C. Results and Discussion

The principal results of the study are presented in Table 2, which shows the means and standard deviations for the number of words recalled correctly and incorrectly for the four conditions of the study, i. e., a crossing-out vowels task (V), a handwriting task (H), and a judgment task with immediate (I) and 48 hour recall (48). The 48 condition, while not directly relevant to the major hypothesis of the study concerning the effect of differential responses, was included to provide some supplementary data on the effect of a moderate time interval between exposure to the stimulus material and its subsequent recall.

TABLE 2

Means and Standard Deviations for the Number of Words Recalled Correctly and Incorrectly for the Four Conditions—Vowels (V), Handwriting (H), Immediate (I), and 48 Hour Delay (48)—of the Study.

		V	H	I	48
Correct	Mean	4.30	8.38	16.86	6.60
	SD	3.00	3.92	7.20	3.20
Incorrect	Mean	14.98	6.34	5.68	12.56
	SD	12.16	5.00	4.22	6.80

It was predicted that as the stimulus material evoked increasing numbers of differential responses, from the V to the H to the I tasks, there should occur a concommitant increase in the recall scores, and inspection of Table 2 indicates such an increase, from a mean of 4.30 words recalled correctly in the V task to a mean of 8.38 words recalled correctly in the H task, to a mean of 16.86 words recalled correctly in the I task. A simple analysis of variance for the V, H, and I tasks yields an $F = 78$, which is highly significant ($P < .001$) and comparisons of V with H, H with I, and V with I tasks are all significant ($P < .01$). The 48 condition shows a mean of 6.60 words recalled correctly, which is significantly different ($P < .01$) from the mean of 16.86 words recalled correctly in the I condition, indicating as might be expected that a delay of 48 hours, from exposure of the stimulus material to its subsequent recall, can produce a marked and significant decrement in recall performance.

With respect to the mean number of words recalled incorrectly, for the V, H, and I tasks, rather than an increase in scores from the V to the I tasks, there is a highly significant decrease (simple analysis of variance yields an $F = 20$, $P < .001$), from a mean of 14.98 words recalled incorrectly in the V task, to a mean of 6.34 words recalled incorrectly in the H task, to a mean of 5.68 words recalled incorrectly in the I task. A comparison of the I and 48 tasks, reveals a significant increase ($P < .01$) from a mean of 5.68 words recalled incorrectly in the I task to a mean of 12.56 words recalled incorrectly in the 48 task. Thus, it appears that: (*a*) as the assumed number of differential responses evoked by the stimulus material increases—from the V to the I tasks—there is a significant increase in the number of words recalled correctly, but a significant decrease in the number of words recalled incorrectly, and (*b*) a delay of 48 hours from exposure to recall, produces a significant decrease in the number of words recalled correctly, in contrast to immediate recall, but a significant increase in the number of words recalled incorrectly.

The results in Table 2 while clearly supporting the main hypothesis of the study, also raise additional issues that require discussion. It would seem that in the type of incidental recall task that is employed in this study, *S* not only produces responses that appeared on the stimulus word-list (i. e., correct recall), but also produces a varying and relatively large number of responses that did not appear on the stimulus word-list (i. e., incorrect recall). If we examine these "incorrect" responses, we find that their number varies systematically as a function of the conditions under which the stimulus word-list was presented (i. e., the V, H, I, and 48 tasks). In general, this variation is what might be expected, in that the number of correct responses is inversely related to the number of incorrect responses. That is, when *S* is unable to recall many correct responses, he apparently elaborates with a large number of incorrect responses, and conversely when *S* is able to recall many correct responses, he greatly reduces the number of elaborated incorrect responses. It almost appears from these results that in the 10 minutes allowed for recall, *S* is set to produce a relatively constant number of total responses, that varies from 14.72 in the H task to 22.54 in the I task, but varying systematically in the proportion of correct and incorrect responses.

To understand further the processes behind the production of incorrect responses, an analysis was made of the content of these responses. In the V condition, which produced a mean of 14.98 incorrect responses, or a total sample for the 50 *S*s of 749 words, there were 601 different words. That is, the production of words was highly unique for each *S*, and in fact only three words (namely, agree, appear, and believe) out of the 601 words produced, were shared as common responses by five *S*s. All other words showing commonality were for fewer than five *S*s. In the H condition, which produced a total sample of 317 words, there were 258 different words, and only one word revealed a commonality as high as five *S*s. For the I and 48 conditions, the results are somewhat different. In the I condition, which produced a total sample of 284 words, there were 170 different words, nine of which showed commonalities of five *S*s or higher. These words, with the commonalities given in parenthesis, were as follows: investment (5), inflation (5), price (6), interest (6), export (8), stock (9), market (9), business (9), and money (11). In the 48 condition, which produced a total sample of 628 words, there were 320 different words, 28 of which showed commonalities of five *S*s or higher, with the highest word "money" showing a commonality of 20 *S*s.

These results suggest that the process governing the production of incorrect responses is not identical in the four conditions of the study. The V and H conditions give rise to large samples of unrelated, unique responses, whose content such as "agree," "appear," etc., bear little or no relation to the concept "economic." The I and 48 conditions, however, are characterized by a higher degree of commonality among the responses and bear a far more direct relationship to the concept "economic," as evidenced by their content such as "investment," "stock," "money," etc. Since the I and 48 conditions involve explicity judging the words for inclusion in the concept "economic," while the V and H conditions do not involve any such judgments, it is not surprising to find more commonality of responses and responses related to the economic concept, in the I and 48 conditions, than in the V and H conditions. It might be assumed then that in the I and 48 conditions, the "economic" judgment task sensitizes *S*'s response hierarchy for "economic" content responses, thus affecting the production of incorrect responses in the manner shown.

If *S*'s response hierarchy for "economic" content-type responses is sensitized by the "economic" judgment task, we should find additional evidence of this in the production of correct responses. Specifically, we might expect that if we examine the content of the correct responses most frequently recalled versus least frequently recalled, there should be systematic differences in the content as a function of the four conditions of the study. Since the I and 48 groups judged the words for inclusion in the economic concepts on a seven-point scale, scale values exist which provide a quantitative estimate of the "economic" content of each word. For example, as shown in Table 1, word number 1, "poem" with a scale value of 1.17 may be said to have low economic content, while word number 100, "industry" with a scale value of 6.32 may be said to have high economic content.

In making the analysis, the mean scale value of the 10 most frequently correctly recalled words and the mean scale value of the 10 most infrequently correctly recalled words were compared. For example, in the I condition the 10 most frequently correctly recalled words (frequency or commonality scores are in parenthesis) were: power (32), boss (29), college (28), imports (22), priest (22), cartel (21), coal (21), society (21), wages (20), and bonds (19). The 10 most infrequently correctly recalled words in the I condition were: comparison (1), drive (1), investigate (1), illumination (1), judgment (1), quality (1), restriction (1), volume (1), violent (1), and civil (1). The mean scale value for the 10 most frequent words given above is 4.94, while the mean scale value for the 10 most infrequent words is 3.50, a difference which is significant at less than the .01 level. For the 48 condition, the corresponding mean scale values are 6.03 and 3.15, a difference which is significant at less than the .001 level. In contrast, for the V and H conditions, the corresponding mean scale values are, for V, 3.50 and 3.99, and for H, 4.86 and 4.74, neither difference yielding statistical significance.

The failure of the V and H conditions to yield a significant difference in scale value between the most frequent and most infrequent recalled words, in contrast to the I and 48 conditions, where the difference was highly significant, indicates clearly the sensitizing effect of the "economic" judgment task to raise the response thresholds for words having high economic content.

That is, in the I and 48 conditions, the words most frequently recalled have high economic content, in contrast to the words most infrequently recalled, whereas in the V and H conditions, where no sensitization to "economic" takes place because of the task instructions, there is no difference in economic content between the words most frequently and most infrequently recalled. The sensitizing effect produced by the "economic" judgment task is evidenced further, when we compare the mean scale values for the most frequently recalled words in the I and 48 conditions. As already indicated these values are 4.94 and 6.03, respectively, a difference which is significant at less than the .05 level. Thus, it appears that a delay of 48 hours from exposure to recall, further magnifies the sensitizing effect, as the most frequently recalled words have an even higher economic content (6.03), than when recall is immediate (4.94). This increased potential to produce responses with high economic content, after a 48 hour delay, probably represents a differential weakening of the response strengths for the words in the list, with the response strengths for the words having high economic content being maintained to a greater degree because of the sensitizing effect of the judgment task.

In the results so far discussed, we have noted both a facilitating and sensitizing effect on recall scores for the "economic" judgment task. A serious methodological question must be raised however, as to the role of the "free production" of economic content words, rather than their recall. That is, in the recall test, in both the I and 48 conditions, to what extent is *S* "recalling" words previously exposed to him, and to what extent is *S* "freely producing" words he believes possess economic content? In an attempt to answer this question, an additional group of 25 *S*s were asked to produce, during a 10-minute period, all the words they felt might be included in the concept "economic." For this sample of 25 *S*s, the mean number of words produced was 58.20, yielding a total sample of 1455 words, of which 702 were unique responses. A comparison then between the mean number of words produced in 10 minutes (58.20) versus the mean number of words recalled in 10 minutes (16.86), indicates a very significant difference ($P < .001$) between the two conditions. These results imply that if there is any tendency to "free produce" in the recall situation, it represents a drastically reduced type of free production, in contrast to the pure production condition.

Another type of comparison between the free production and recall situations may be made with regard to the commonality of responses in the two situations. If we compare the 20 most frequently produced words in the free production situation with the 20 most frequently recalled words in the I and 48 conditions, we find for the 48 condition only four of the 20 words identical in both samples, and for the I condition only three of the 20 words identical in both samples. These results, that is, both total number of responses produced and commonality of responses, are suggestive that the processes underlying response production in the free production and recall situations are different, and therefore, in the recall situation, the facilitating and sensitizing effects of the "economic" judgment task are not explainable solely in terms of a tendency of *S*s to merely free produce words having economic content, but some active recall process also must be involved.

D. SUMMARY

Since Postman has suggested that incidental learning is a positive function of the number of differential responses evoked by the stimulus material, this study was designed to manipulate the number of such differential responses by varying the operations to be performed on the stimulus by *S*. Specifically, *S*s were presented with a list of words and instructed to perform certain operations on the words, i. e., one group was instructed to cross-out all vowels, another group to copy the list, and a third group to judge the words for inclusion in the concept "economic." Following completion of these tasks, which are assumed to evoke an increasing number of differential responses, *S*s were tested for their retention of the word-list. The results showed clearly that as the number of differential responses increased, from the vowel to the copying to the judgment tasks, there was a highly significant increase in the recall scores, as expected.

REFERENCES

1. McGeoch, J. A., & Irion, A. L. The Psychology of Human Learning. New York: Longmans, Green, 1952.
2. Myers, G. C. A. A study in incidental memory. *Arch. of Psychol.*, 1913, **4**, No. 26.
3. Postman, L., & Phillips, L. W. Studies in incidental learning: I. The effects of crowding and isolation. *J. Exp. Psychol.*, 1954, **48**, 48-56.
4. Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and of the method of testing. *J. Exp. Psychol.*, 1955, **49**, 1-10.
5. Postman, L., & Adams, P. A. Studies in incidental learning: VIII. The effects of contextual determination. *J. Exp. Psychol.*, 1960, **59**, 153-164.
6. Pritchard, L., Mayzner, M. S., & Tresselt, M. E. The Mayzner-Tresselt word-concept check-list. *Psychol. Newsl.*, 1955, **7**, 27-33.

15 Sydney Avenue
Deal, New Jersey

Published as a separate and in *The Journal of Psychology,* 1960, **50**, 349-354.

SWITCHING BEHAVIOR IN A PROBLEM SOLVING TASK*

New York University and Barnard College of Columbia University

M. E. TRESSELT AND M. S. MAYZNER[1]

A. PROBLEM

Typically in very complex problem solving tasks, such as those encountered in science and engineering, more than one approach to a problem may initially seem attractive. In fact, several alternatives may be present in the early stages, and one of the first tasks confronting the problem-solver is to adopt some strategy for testing or exploring the various alternatives. For example, in developing some new chemical process, a wide range of catalysts may be available, only one or a few of which may possess the properties necessary for a successful reaction to occur. Therefore, the chemist must develop some method or approach in order to choose a particular catalyst to begin his investigation with, and further, he must decide when to switch to another catalyst, if the initial choice appears not to have been the correct one.

The present investigation examines the effects of two variables, in a multi-alternative problem solving task, on the total number of switches occurring in a given period of time. The two variables are: (*a*) the probability that any given alternative will lead to a solution, and (*b*) the total number of alternatives present. It is assumed that if all the alternatives present have unity probability of leading to a solution, then the number of switches occurring in a given period of time should be different than if only one of the alternatives present has unity probability while the remaining alternatives have zero probability of leading to a solution, with the former condition producing a fewer number of switches than the latter. What is critical, however, in producing differences in the number of switches occurring, is *S*'s estimate or expectation of the probability of a solution associated with any given alternative, rather than the "actual" probabilities. Further, the number of switches occurring should vary as a function of the number of alternatives present, with more alternatives evoking a greater number of switches.

---

*Received in the Editorial Office on August 27, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to express their appreciation to Dr. H. M. Jenkins of Bell Telephone Laboratories who collaborated with them in the initial formulations of this study.

B. Method

1. *Subjects*

All *S*s were enrolled in undergraduate psychology courses at New York University. Forty *S*s were employed in Experiment I and 80 *S*s were employed in Experiment II.

2. *Materials and Procedure*

In attempting to construct a problem solving task in which the number of alternatives to reach a solution could be precisely controlled and the number of switches made by *S* in changing from one alternative to another could be precisely measured, the following was adopted. The *S* was presented with a set of five-letter anagrams, each typed separately on a 3 × 5 card and placed face down, in a row before him. The *S* was instructed that before him were a set of cards and on each card a single anagram appeared. He was told his task was to obtain a single anagram solution, which could be accomplished by solving any one of the set of anagrams before him, but he could only work on a single anagram at a time, the others remaining face down. However, he could switch back and forth from one anagram or alternative to another as often as he wished.

In Experiment I, four different conditions, with 10 *S*s per condition, were examined. In all four conditions, *S* was presented with four anagrams face down, and instructed that the solution of any of these constituted a solution of the problem. In the first condition, *S* was told that all four anagrams had solutions—the 4S condition; in the second condition, *S* was told that only three of the four anagrams were solvable, the remaining one was not—the 3S condition; in the third condition, *S* was told that only two of the four anagrams were solvable, the remaining two were not—the 2S condition; in the fourth condition, *S* was told that only one of the four anagrams was solvable, the remaining three were not—the 1S condition. Thus, the instructions for these four conditions were designed to produce differences in *S*'s estimate or expectation of the probability of a particular anagram being solvable or not solvable. For example, in the 4S condition, *S*'s estimate should be unity probability for solution for all four anagrams, while in the 1S condition, *S*'s estimate should be unity probability for solution for one of the four anagrams and zero probability for solution for the remaining three anagrams. Since *S*, of course, does not know which one of the four anagrams is solvable and which three are not, one might expect many more switches for the 1S condition, where *S* is attempting to find the one solvable anagram or alternative, than for the 4S condition, where as far as *S* is concerned one anagram or alternative is as good as another for reaching a solution to the problem.

In Experiment II, a two by four factorial design was employed, with 10 *S*s per cell, or a total of 80 *S*s. One variable involved the number of anagrams or alternatives presented to *S*. Thus, *S* was presented with either two, six, eight, or 10 anagrams, face down before him. The other variable involved the instructions, with half of the *S*s being told that all of the anagrams were solvable—the AS condition; while the remaining half were told that only one of the set facing them (i.e., two, six, eight, or 10 anagrams) was solvable—the 1S condition.

In both Experiments I and II, all *S*s were tested individually and assigned to the various experimental conditions at random. Each *S* received only one problem and was allowed eight minutes to achieve a solution. The dependent variable was the number of switches or changes from one anagram to another, each *S* made during the eight minute period. To insure that the full eight-minute period would be constant for each *S*, none of the anagrams in any of the conditions had solutions. None of the *S*s expressed any suspicion of this, since each *S* was presented with only a single problem and they were instructed that the anagrams were all relatively difficult.

C. Results and Discussion

The results of Experiment I are presented in Table 1, which shows the mean number of switches for the four experimental conditions, i.e., 4S 3S, 2S, and 1S. A simple analysis of variance indicates that the progressive increase in the mean number of switches from 7.3 for the 4S condition to 13.1 for the 2S and 1S conditions is significant at less than the .01 level. These findings generally are in line with our expectations. That is, as S's estimate or expectation of the probabilities for solution of a set of alternatives goes from unity for all alternatives to unity for only one alternative and zero for the remaining, there will be an increase in the

TABLE 1
Means and standard deviations for the number of switches for the four conditions of Experiment I, i.e., all anagrams solvable (4S), three anagrams solvable (3S), two anagrams solvable (2S), and one anagram solvable (1S).

	Conditions			
	4S	3S	2S	1S
Mean	7.3	7.9	13.1	13.1
SD	4.2	2.6	5.6	5.1

number of switches as *S* attempts to locate that alternative which will lead to problem-solution. The relationship, however, between the number of switches and *S*'s estimates or expectations that a particular array of alternatives are all equally certain or not certain of yielding a solution, is not a linear one. There is, in fact, essentially no difference in the mean number of switches between the 4S and 3S conditions and between the 2S and 1S conditions, with the main difference occurring between the 3S and 2S conditions. This would suggest that a particular level or balance between the number of certain and uncertain alternatives must be reached before an appreciable increase in the number of switches will occur.

The findings of Experiment II are presented in Table 2, which shows the mean number of switches for the various experimental conditions. An analysis of variance (a square-root transformation was performed to obtain greater homogeneity of variance) indicates that both main effects, number of alternatives (two, six, eight, or 10) and number of solvable anagrams (all or one), are highly significant ($P < .001$). The interaction term is not significant. As was expected, as the number of alternatives is increased from two to 10, there is a corresponding increase in the mean number of switches occurring for both the AS and 1S conditions. For the AS condition, for the two, four, six, eight, and 10 alternative cases, we have 3.2, 7.3, 7.3, 7.3, and 7.9 switches, respectively. The values for the four alternative case are included here also, as obtained from Experiment I. For the 1S condition, for the two, four, six, eight, and 10 alternative cases, we have 6.7, 13.1, 17.6, 14.2, and 16.7 switches, respectively.

TABLE 2

Means and standard deviations for the number of switches for the eight conditions of Experiment II, i.e., two, six, eight, and ten alternatives with all anagrams solvable (AS) and with only one anagram solvable (1S).

		Alternatives			
		2	6	8	10
AS	Mean	3.2	7.3	7.3	7.9
	SD	1.2	4.5	3.3	5.3
1S	Mean	6.7	17.6	14.2	16.7
	SD	2.2	4.8	5.3	7.1

The increase in the number of switches which occurs as we increase the number of alternatives probably reflects a strategy designed to uncover as quickly as possible that anagram or alternative which *S* can most readily solve in the AS condition or the one single anagram or alternative which can be solved in the 1S condition. However, this increase in the number of switches certainly is not a linear function of the number of alternatives present, and the results strongly suggest that the only appreciable increase in the number of switches occurs when the number of alternatives increases from two to four. After this there is a leveling off of the function for both the AS and 1S conditions.

We find again in Experiment II, a highly significant increase in the number of switches between the AS and 1S conditions, which is equivalent to the significant increase in Experiment I between the 4S and 1S conditions, which is relatively independent of the number of alternatives present. Actually, one might have expected to find differences between the AS and 1S conditions to increase consistently as the number of alternatives increases, since *S* for the 1S condition in contrast to the AS condition must search an increasing number of alternatives to find the single one having a solution. In fact, however, the differences between the AS and 1S conditions for the six, eight, and 10 alternative cases are quite similar. Apparently the strategy adopted then is not to search at an increasingly rapid rate in the 1S condition as the number of alternatives is increased, but to level off at about 15 switches for an eight-minute period, and to level off at about 7 switches for an eight-minute period in the AS condition.

In addition to the variables examined here which may affect switching behavior, the following factors deserve consideration. First, in the attempt to precisely control the time spent on the problem, none of the alternatives were solvable. By introducing alternatives which have solutions, while it is true there is a loss of certain controls, there could be a gain in the ability to examine the effects of various patterns of successful problem-solutions on subsequent switching behavior. Second, a far greater range of probability values associated with the alternatives should be investigated. Specifically, *S*s could be treated to various "probability mixes" of the alternatives present. That is, with four alternatives, *S*s might be informed that one alternative has a unity probability of solution, another alternative has a .50 probability of solution, while the remaining two alternatives have a zero probability of solution, or they might be informed that each of the four alternatives present have a .25 probability of solution or a .75 probability of solution. Further, analysis of the time spent on each alternative as well as the number of switches occurring should prove instructive.

D. SUMMARY

The present investigation was designed to examine the effects of two variables in a multi-alternative problem solving task on the number of

switches occurring in a given period of time. The two variables were: (*a*) the probability of any given alternative yielding a solution, and (*b*) the number of alternatives present. The results indicated that an increase in the number of switches occurs as a function, which is non-linear, of both an increase in the number of alternatives having zero probability of solution and in the total number of alternatives present.

15 Sydney Avenue
Deal, New Jersey

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 355-376.

THE SIGNIFICANCE OF DYSPHASIA FOR INTELLIGENCE AND ADAPTIVE ABILITIES*

Department of Neurology, Indiana University Medical Center

RALPH M. REITAN[1]

A. INTRODUCTION

The relationship between language and thinking has been studied by investigators from many scientific disciplines during the last century. The problem has also been considered by philosophers and poets not only recently but in the earliest recorded writings. An excellent review of the background of the problem, published by Meyers (6), describes the continuing historical interest that identifies this problem as one that has persistently concerned man in his attempt to understand himself.

John B. Watson, generally regarded as the father of behavioristic psychology, proposed the most intrinsic of relationships between language and thinking—that thinking is subvocal speech (17). His influence is quite pronounced in the writings on this topic of most present-day psychologists although few would accept the implication that thinking is impossible without speech. Nevertheless, one frequently encounters statements such as: "Language is inseparably bound up with thinking and other symbolic behavior (4), and " . . . we must admit that the symbols which represent most of the world are language symbols (verbal, gestural, or written) and that most of our thinking appears to be an internal manipulation of such symbols" (7).

Meyers has pointed out that investigation of dysphasia offers " . . . a singular opportunity to inquire into the language-thought relationship" (5, p. 65). In the neurologic literature, however, the question has been stated principally with respect to the integrity of language functions in relation to intelligence. Recent reviews (8, 3) have indicated the diversity of opinion concerning the possible reciprocal significance of aphasia and intellectual impairment. A comment by Weisenburg and McBride (19) summarizes the problem:

---

* Received in the Editorial Office on August 30, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was aided by a grant from the United Cerebral Palsy Research and Educational Foundation. The writer wishes to acknowledge the assistance of Miss Phyllis Bramer with the statistical computations.

> The extent to which intelligence may be said to be affected is a moot point. Some investigators have held that intelligence suffers because speech is disturbed while others have held that speech is disturbed because of the intellectual disorder.

The only controlled experimental investigations of this problem using dysphasic subjects are those of Meyers (6, 5) and Reitan (8, 9, 10). Meyers first used such subjects in this connection in 1940 and should be credited with devising the experimental technique and suggesting elaboration of the method for future experimentation. Meyers obtained results suggesting that his dysphasic subjects suffered no disadvantage in solving a non-verbal abstraction type of problem as compared to nondysphasic control patients. He declined to generalize conclusions from the study, however, for a number of reasons including small sample-size and insufficient variability in the type of problem-solving procedure used. Reitan also used a small sample, but the testing material differed from Meyers' study in that the Rorschach Test and Wechsler Mental Ability Scale were used for deriving the dependent variables. Reitan's evidence also failed to show that the dysphasic subjects experienced any special handicap in these testing situations, except on tests that required rather specifically oriented verbal responses (the verbal subtests of the Weschsler Mental Ability Scale). The basis for generalizing from these experimental results is strengthened by the fact that two different investigators, using completely discrete groups of dysphasic and controlled patients, found results of very similar significance even though different testing procedures were used. Nevertheless, the variety of testing procedures was still limited and the groups used by each investigator were small.

B. Hypotheses

The present investigation compared groups (somewhat larger than those previously used) of very carefully selected patients who were individually examined with an extensive battery of psychological tests. The study was designed: (*a*) to provide additional evidence regarding the validity of psychological tests with respect to measurement of the effects of brain damage, and (*b*) to compare the performances on a wide variety of psychological tests of brain-damaged patients with and without dysphasia. We anticipated that such comparisons should provide a basis for inferring in greater detail than heretofore possible the particular significance of dysphasia with respect to more general psychological functions and to contribute experimental evidence relevant to the problem of relationships between language and thinking.

C. SUBJECTS

Three groups of 32 subjects were used. One group was composed of patients with brain lesions and dysphasia, another of patients with brain lesions but without dysphasia, and the third of patients hospitalized for various reasons but with no anamnestic or clinical evidence of organic brain dysfunction. A matched-pairs design was used in which each subject was matched with corresponding subjects in the other groups for sex, color, chronological age, years of formal education, and type of brain lesion. Selection of 32 patients for each group in accordance with these criteria required a review of medical records and results of aphasia examinations for more than 1,500 patients. All patients selected were right-handed, white, and cooperative during the testing. The Halstead-Wepman Aphasia Screening Test had been administered to all patients considered for inclusion in the study in a standard manner and with detailed recording of all responses. This test was administered at the same time as the additional psychological testing so that we could be sure that the presence or absence of dysphasia was concurrent with the test results used as dependent variables. Table 1 presents certain summary statistics describing the groups.

TABLE 1

SUMMARY STATISTICS FOR GROUPS OF DYSPHASIC, NONDYSPHASIC, AND CONTROL SUBJECTS WITH RESPECT TO SEX DISTRIBUTION, AGE, AND EDUCATION

	N	Male	Female	Mean age	SD	Mean education	SD
Dysphasic	32	30	2	37.00	11.23	10.47	3.04
Nondysphasic	32	30	2	36.69	10.38	10.63	2.98
Control	32	30	2	35.47	10.78	10.44	3.33

The control subjects were selected from a group for whom neither the clinical neurologic examination nor anamnestic information had yielded positive or even suggestive findings with respect to organic impairment of brain functions. The diagnostic categories for these patients were determined from detailed examinations during hospitalization and included: neurosis, 18; schizophrenia, 4; paraplegia, 6; and surgical lesions of various organs not involving the central nervous system, 4. No particular attempt was made to structure the diagnostic distribution of the control group other than to adhere as rigorously as possible to the criterion of normal organic brain functions. However, we did feel that it was important to use patients in composing this group rather than normally functioning persons in order to impose at least roughly an equation in all groups of the anxieties, changes in normal routine, and other factors that may attend hospitalization.

Our most difficult problem occurred in the attempt to compose the two brain-damaged groups (dysphasic and nondysphasic) with respect to matching the subjects in pairs having lesions of the same type. Table 2 presents the diagnostic distributions for these groups.

TABLE 2
DIAGNOSTIC DISTRIBUTIONS IN TWO BRAIN-DAMAGED GROUPS

Diagnosis	Dysphasics	Nondysphasics
Glioblastoma multiforme	5	2
Astrocytoma	1	3
Oligodendroglioma	1	0
Ependymoma	0	1
Metastatic melanoma	0	1
Meningioma	1	1
Arteriovenous malformation	1	1
Cerebral occlusion	5	4
Cerebral hemorrhage	1	1
Intracerebral hematoma	2	3
Subdural hematoma	1	1
Closed head injury	6	7
Depressed skull fracture	3	2
Temporal lobe epilepsy	2	2
Cerebral abscess	1	1
Multiple sclerosis	2	2
	32	32

No controlled comparisons of possible differential effects of various types of brain lesions on psychological test results have appeared in the literature to our knowledge. The differential significance of certain types of lesions is well known to neurologists, however, and we feel confident that they affect psychological functions differentially as well. Preliminary evidence (14) indicates that the type of lesion can be inferred in individual cases from the psychological tests used in this study with accuracy that far exceeds chance expectancy. These considerations led us to make every effort to match the patients at least approximately with respect to type of lesion, since we felt that this procedure would probably be quite important to minimize uncontrolled variance and thus enhance the probability of obtaining results more specifically pertinent to the hypotheses. We did not achieve this matching with perfect success. Six of the 32 matched pairs did not have exactly the same types of brain lesions, as shown in Table 3.

We felt that inclusion of these six pairs to increase the size of the groups would add more than it would detract from the value of the study since the misclassifications occurred within rather than between more general categories. In all of the other 26 pairs of subjects the diagnosis was exactly the same for each subject.

Meyers (6, 5) apparently feels that dysphasic patients with receptive disorders represent the best group to use in this type of study. Expressive language disorders, however, seem equally pertinent to our hypotheses since they are frequently present as a part of the dysphasic picture. Careful review of the symptoms shown by the patients in our dysphasic group indicated that 10 had predominantly expressive symptoms, 18 had prominent symptoms of both expressive and receptive types, and four had predominantly receptive dysphasia. The group had a mean of 5.12 symptoms (SD, 2.37) that were distinctly elicited. Finger dysgnosia, dystereognosis, and construction dyspraxia were also tested routinely, but were infrequently found in this group. The frequency with which various symptoms occurred is indicated in Table 4.

TABLE 3
PATIENTS MISMATCHED WITH RESPECT TO TYPE OF BRAIN LESIONS IN THE DYSPHASIC AND NONDYSPHASIC GROUPS

Patient-pair number	Dysphasic group	Nondysphasic group
2	Glioblastoma multiforme	Ependymoma
3	Glioblastoma multiforme	Astrocytoma
5	Oligodendroglioma	Astrocytoma
7	Glioblastoma multiforme	Metastatic melanoma
14	Cerebral occlusion	Intracerebral hematoma
31	Depressed skull fracture	Closed head injury (linear skull fracture)

The Halstead-Wepman Aphasia Test had been administered to the other

TABLE 4
FREQUENCY OF OCCURRENCE OF VARIOUS DYSPHASIC AND RELATED SYMPTOMS IN THE EXPERIMENTAL GROUP OF 32 DYSPHASIC PATIENTS

Symptom	Number	Percentage
Dysnomia	21	66
Dysgraphia	21	66
Dyslexia	20	62
Dyscalculia	19	59
Spelling dyspraxia	23	72
Central dysarthria	22	69
Right-left disorientation	12	38
Visual letter dysgnosia	10	31
Auditory verbal dysgnosia	9	28
Visual number dysgnosia	4	12
Body dysgnosia	3	9
Finger dysgnosia—right hand	9	28
Finger dysgnosia—left hand	3	9
Dystereognosis—right hand	4	12
Dystereognosis—left hand	1	3
Construction dyspraxia	4	12

two groups also, but, in accordance with criteria for their composition, none of the patients included had any symptoms of dysphasia.

D. PROCEDURE

Halstead's tests of biological intelligence (2, 11), the Trail Making Test (12, 13, 15), and the Wechsler-Bellevue Scale of Adult Intelligence (18) were individually administered to each subject by technicians who were completely unfamiliar with diagnostic information regarding any patient. Final scoring of tests for each subject was completed before the groups were composed. In order to facilitate understanding of the intergroup differences, a description of the psychological tests will be presented together with the results.

The first step in the analysis of the data was to combine results for the three groups on each measure and convert the raw score distributions to ranked T-score distributions. This procedure was followed because a variety of differently scaled psychological tests were being used for which variability may have differed in the three groups. Secondly, conversion of raw score distributions to normalized distributions minimized possible difficulties in appropriate statistical analysis with parametric techniques. A T-score distribution with a mean of 50 and a standard deviation of 10 was obtained for the combined groups for each test. The conversion was done in such a way that the highest T-score corresponded with the best performance. The individual subjects were then returned to their appropriate groups and means based on T-scores were obtained for each variable. Next, T-score differences between matched pairs were obtained for each pair of groups. Intergroup comparisons of the means of these difference-score distributions were then made in order to determine the probability that the mean difference scores were attributable to the operation of chance factors.

E. RESULTS

Figure 1 presents graphically the results obtained with Halstead's Neuropsychological Battery and the Trail Making Test.

Although the results will be examined for each test individually, one can see at a glance that the control group performed much better on every test than did either of the groups with brain damage. In general, the performances of the two brain-damaged groups were approximately equivalent, suggesting that dysphasia *per se* apparently has only certain specific effects on the results.

Figure 2 presents results for the three groups on the Wechsler-Bellevue Scale using the same ordinate as that used in Figure 1.

The distribution of groups is somewhat different than in Figure 1, particularly with respect to the verbal tests. Differences between the control group and the brain-damaged group without dysphasia failed to show the consistently significant differences that were present with Halstead's tests and the Trail Making Test. The dysphasic group, however, did very poorly on the verbal test in comparison with the other two groups. The performance tests were done at a lower level than the verbal tests by the nondysphasic group but at a higher level than the verbal tests by the dysphasic group. Thus the performance tests, which in appearance are limited very little by verbal functions, gave results that were almost identical for the two brain-damaged groups.

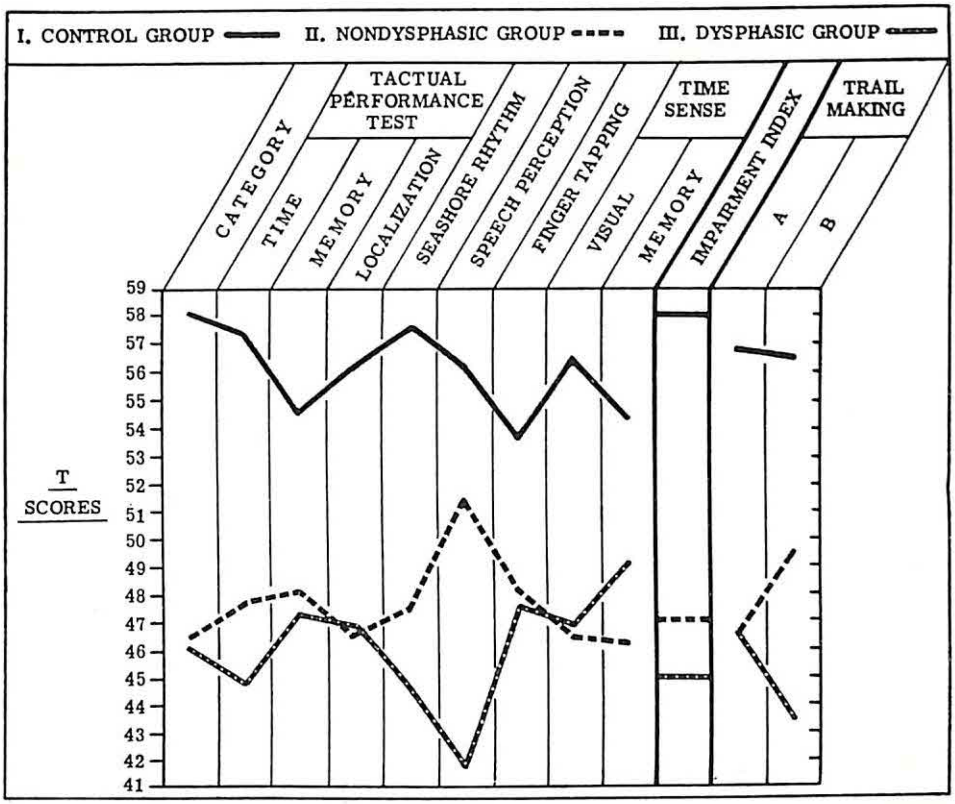

FIGURE 1

Graphic Representation of Results Obtained by Dysphasic, Nondysphasic, and Control Groups on the Halstead Neuropsychological Test Battery and the Trail Making Test

Additional understanding of the effects of dysphasia on a variety of tasks requiring adaptive intelligence and problem-solving ability can be derived from consideration of the results on the individual tests.

1. *The Halstead Neuropsychological Battery*

a. Category Test. This test utilizes a projection apparatus for presentation of 208 stimulus figures on a milk-glass screen in front of the patient, who is told that each figure will suggest a number from 1 through 4. He is required to select one of four levers, which are numbered from 1 to 4, as his response to each stimulus figure. Depression of any lever will cause either a bell or buzzer to sound depending upon whether the lever selected is the "right" or "wrong" answer. Only one response is allowed for each item. The subject is told that the stimulus figures are divided into seven groups, and that each group has a single principle running through the entire group. On the first item in a group he can only guess, but as he progresses through the items of the group the bell or buzzer accompanying

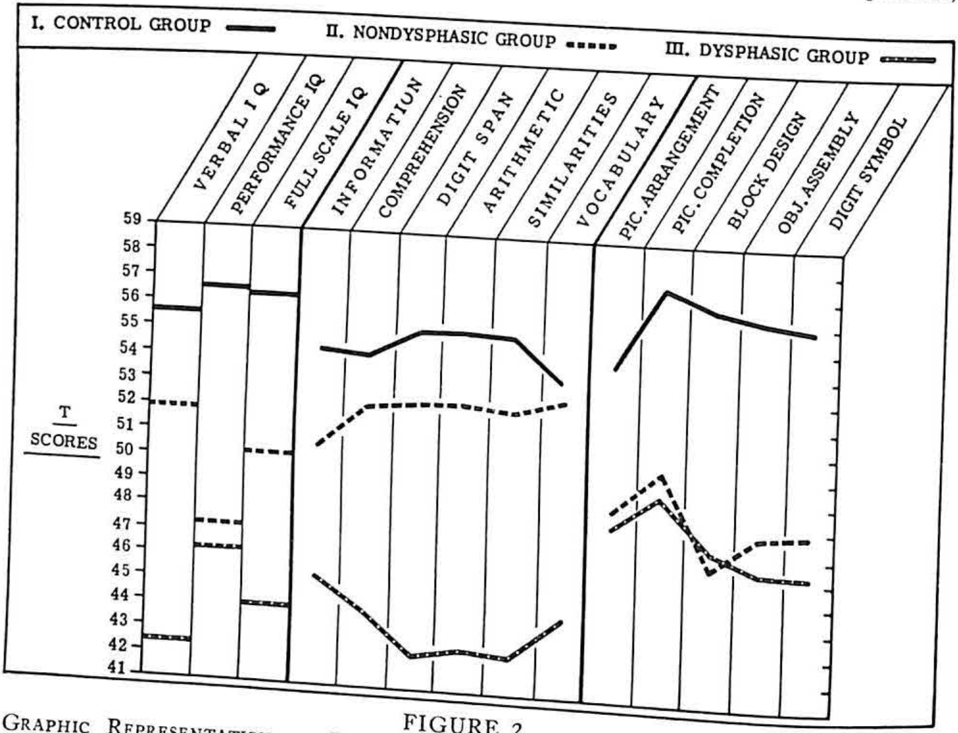

FIGURE 2

Graphic Representation of Results Obtained by Dysphasic, Nondysphasic, and Control Groups on the Wechsler-Bellevue Scale of Adult Intelligence (Form I) each response will permit testing of one possible principle after another until an hypothesis is hit upon which is positively reinforced consistently by the bell. The subject is never told the principle for any group regardless of the difficulty he encounters, but the first and second groups are nearly always easily performed. The first group presents a Roman numeral for each stimulus figure and the bell ("correct" response) sounds if the subject depresses the lever having the same number as the one appearing on the screen. In the second group the subject must learn to press the lever whose number corresponds to the number of items appearing on the screen. Three identical squares might appear on the screen as an item early in the group (for which the lever numbered "3" would be correct) but the content changes as the items progress. Later in the group three different kinds of figures might appear as an item, but since they are still in the second group and the principle is the same, the correct answer would still be 3. At the end of a group the examiner routinely says, "That was the last picture in this group. Now we will go on to the next group. The principle might be the same or it might be different. We want you to find out." The third group presents four figures in each item and requires the subject to learn to depress the lever corresponding with the figure which is most different from the others. Although this group begins rather simply, it progresses to items in which one figure may differ from the others in three or more respects (such as size, shape, color, solidness of figure) while the rest of the figures differ from each other in only two respects.

This description probably suffices to indicate the nature of the Category Test, although other principles are used in the remaining groups. There appears to be no doubt that this test is a complex "abstraction" test requiring fairly sophisticated ability in noting similarities and differences in stimulus material, postulating hypotheses that appear reasonable with respect to recurring similarities and differences in the stimulus material, testing of these hypotheses with respect to reality considerations (the bell and the buzzer), and learning through adaptation of the hypotheses in accordance with the positive or negative reinforcement accompanying each response. It would certainly seem that this test requires thinking ability and perhaps even thinking ability of quite a high order.

The interesting aspect of our results on the Category Test is not the striking difference between the control group and the groups with brain damage, since highly reliable differences in such comparisons have been presented before (2, 11). In fact, previous studies suggest that the Category Test is one of the most sensitive psychological tests to the effects of cerebral dysfunction that has ever been devised. The interesting finding of this study relates to the comparative performances of the brain-damaged groups with and without dysphasia. One might certainly expect that the dysphasic group (in which the majority of patients suffered from clinically recognized organic impairment of abilities in reading, writing, calculating, *and* naming of common objects) would be handicapped since implicit verbalizations would appear to contribute to a good score on the Category Test. Nevertheless, the groups obtained nearly identical results. The dysphasic subjects appeared to have no special handicap in comparison with the brain-damaged group without dysphasia. Our results suggest that the abilities the Category Test measures are seriously impaired by organic brain damage, but that the presence or absence of dysphasia *per se* it not specifically relevant to the results obtained.

b. Tactual Performance Test (Time, Memory, and Localization components). This test utilizes a modification of the Seguin-Goddard formboard. The subject is blindfolded before the test begins and is not permitted to see the formboard or blocks either before or after the test. His task is first to fit the blocks into their proper spaces on the board with his preferred hand. Next, he performs the task again using his other hand. Finally, he is asked to do the task a third time using both hands. The Time component of the test is based on the total time needed to complete these three trials. After the board and blocks have been put out of sight, the blindfold is removed and the subject is required to draw as well as possible a diagram of the board with the blocks placed in their correct positions. The Memory component of this test is based upon the number of shapes correctly reproduced in the drawing of the board, and the Localization component is based on the number of blocks correctly localized.

The Tactual Performance Test is another complex test. Placing of the variously shaped blocks in their proper spaces on the board depends upon tactile form discrimination, kinesthesis, coordination in movement of the upper extremities, manual dexterity, and the ability to visualize the spatial configurations of the shapes in terms of their interrelationships.

The control group performed much better on each of the three measures derived from this test than did either brain-damaged group (Figure 1). The dysphasic group probably required more time to place the blocks than did the nondysphasic group ($p < .05$), but no differences were present on the Memory and Localization scores from the drawings. Again, one might easily have postulated that dysphasia would be a distinct disadvantage in this performance because of impairment in implicit naming of the shapes

(66 per cent of the dysphasic patients showed evidence of dysnomia), but the experimental results imply that the handicaps experienced by the nondysphasic subjects were almost equally disadvantageous with respect to a successful performance.

c. Seashore Rhythm Test. This is a subtest of the Seashore Tests of Musical Talent. The subject is required to discriminate between 30 pairs of rhythmic beats that are of the same or different rhythms. This test requires alertness, sustained attention to the task, and the ability to perceive differing rhythmic sequences. In agreement with previous studies (2, 11, 16), highly reliable differences were present in comparing the mean performances of the groups with and without brain damage. Although the dysphasics obtained a somewhat lower mean score than did the nondysphasics, the difference in the groups was not sufficiently consistent to permit a conclusion that factors other than chance variability were responsible ($p < .20$).

d. Speech-sounds Perception Test. This test consists of 60 spoken nonsense words which are variants of the "ee" sound. The test is played from a tape recorder with the intensity of sound adjusted to the subject's preference. The subject's task is to select the spoken syllable from four alternatives printed for each item on his test form. This is the only test in Halstead's battery that deals specifically with verbal material and the only test on which the dysphasic patients clearly performed more poorly than did the nondysphasic patients. This result was not surprising in consideration of the requirements of the test. In addition to sustaining attention through the 60 items, the subject must perceive the spoken stimulus-sounds through hearing and relate these perceptions through vision to the correct configuration of letters on the test form. Auditory verbal dysgnosia was present in 28 per cent of the dysphasic group and dyslexia was present in 62 per cent. In the other tests one might postulate that language functions would be advantageous ancillary mechanisms, but in the Speech-sounds Perception Test language functions were required specifically.

e. Finger Tapping Test. This test is a measure of finger tapping speed with the index finger of the preferred hand. The subject is given five separate 10-second trials with the hand held in a constant position in order to be sure to require movements of only the finger rather than the whole hand and arm. Every effort is made to encourage the subject to tap as fast as he possibly can. Although this test yielded a highly significant difference between equated groups with and without brain damage in a previous study (11), the present results are of marginal significance. Our results have shown consistently that the Finger Tapping Test is not as reliably affected by brain damage as is the Tactual Performance Test (11, 14) (described above), although both are dependent to a large extent upon motor functions. The Finger Tapping Test is more a measure of pure motor speed, however, while the Tactual Performance Test is much more complex in nature. On the basis of these and other findings one might propose the hypothesis that the effects of brain damage are frequently complex rather than isolated or discrete, and that a complex task is better able consistently to demonstrate the effects of brain damage than is a more pure type of measure.

f. Time Sense Test (Visual and Memory Components). This test requires the subject to depress a key which permits a sweep-hand to rotate on the face of a clock. The subject's task is to allow the hand to rotate 10 times and then to stop it as close to the starting (vertical) position as possible. The Visual component of this test is scored as the amount of error in 40 trials. The Memory component is the amount of error on 20 trials that are interspersed in series of 10 after the first 20 visual trials have been completed. The face of the clock is turned away during the Memory trials and the subject is asked to duplicate the visually controlled performance as closely as possible.

The Visual component of this test requires the subject to maintain alertness and coordinate counting from one to 10 with the rotations of the clock's sweep-hand. Visuo-motor coordination is also required in order to stop the hand's rotation in the correct position. The Memory component requires estimation of the duration of time necessary for the hand to make 10 revolutions. The measures derived from this test showed reliable differentiation of the brain-damaged groups from the control group, but there was not the least evidence to suggest a reliable difference between the dysphasic and nondysphasic groups.

g. Halstead Impairment Index. The Impairment Index is a composite value based upon the 10 tests in the Halstead battery and is determined for an individual subject merely by counting the number of tests on which the results fall into the range characteristic of the performance of brain-damaged rather than normal subjects. The Impairment Index was based on eight rather than all 10 of Halstead's neuropsychological measures in this study because our data was incomplete with respect to the two measures based on critical flicker frequency. The estimated Impairment Index, however, showed highly significant differentiation between the control and brain-damaged groups, providing another verification of the sensitivity of this measure to organic brain dysfunction. Again, the dysphasic and nondysphasic groups did not differ significantly.

2. *Trail Making Test*

The Trail Making Test consists of two parts, *A* and *B*. Part *A* is composed of 25 circles distributed over a white sheet of paper that enclose numbers from 1 through 25. The subject is required to connect the circles with a pencil line as quickly as possible, beginning with 1 and proceeding in numerical sequence to 25. Part *B* consists of 25 circles numbered from 1 through 13 and lettered from *A* through *L*. The subject is required to connect the circles, alternating between numbers and letters as he proceeds in ascending sequence. The score is obtained as the number of seconds required to finish each part.

The results shown in Figure 1 substantiate previous indications of the sensitivity of the Trail Making Test to organic brain damage (12, 13, 15, 1) since the controls performed better than either brain-damaged group at highly significant levels of confidence. Even though 12 per cent of the dysphasic group gave evidence of visual number agnosia on the aphasia examination, this group performed exactly as well on Part *A* as did the nondysphasic group. Part *B* probably places a more stringent requirement on the subject in use of symbols than does Part *A*, however, and the dysphasic subjects performed significantly more poorly on Part *B* than did the nondysphasic group. This test and the Speech-sounds Perception Test were the only ones so far considered that indicated a performance by the dysphasic group which was clearly poorer than that of the nondysphasic group. It is interesting to note that these tests also were the only ones to include language symbols as intrinsic parts of the tests.

3. *Wechsler-Bellevue Scale*

a. Verbal Subtests. These tests (Information, Comprehension, Digit Span, Arithmetic, Similarities, and Vocabulary) are each dependent upon specific and standardized verbal stimuli presented by the examiner and upon verbal responses by the subject. Specific illustrations of their content include procedures such as asking the subject to name the capital of Italy, to explain why laws are necessary, to repeat digits after the examiner, to perform mental arithmetic, to indicate similarities between such things as a poem and a statue, and to define words. While tests of this type long have been considered as valuable indicators of intelligence, obviously successful performance is heavily dependent upon adequate verbal functions.

The outstanding factor in the intergroup comparisons (Figure 2) was the poor performance of the dysphasic subjects. The nondysphasic group, even though composed of patients with verified brain damage, was much more similar to the controls than it was to the dysphasic group. The results suggest that these tests are relatively insentitive to the effects of brain damage unless dysphasia is present. With the presence of dysphasia, however, one can hardly be surprised that verbal tests of this nature reflect a poor performance.

b. Performance Subtests. The performance subtests of the Wechsler-Bellevue Scale include Picture Arrangement, Picture Completion, Block Design, Object Assembly, and Digit Symbol. The subject is required to perform tasks such as to place disarranged pictures in the proper sequence to tell a meaningful story, to identify missing parts in a series of pictures, to arrange colored blocks to reproduce the configurational pattern shown on a picture, to place disarranged parts in the proper spatial configuration to form objects such as a manikin, and to substitute symbols for numbers as quickly as possible.

The results on these tests indicated consistently that the two brain-damaged groups performed more poorly than did the controls. Again, however, the dysphasic and nondysphasic groups obtained nearly equivalent scores. None of the differences between the two brain-damaged groups approached statistical significance.

4. *Illustrative Results on Individual Patients*

Although statistical comparisons of matched groups provide a preferable basis for scientific conclusions, the results of this study can be supplemented meaningfully by illustrations with individual patients. (The writer is indebted to Dr. Robert F. Heimburger, Director of Neurological Surgery, Indiana University Medical Center, for reviewing the medical records and providing the information on which these summaries are based.)

a. Patient C. H. This 30-year-old, right-handed, white male was admitted on April 26, 1956, with a history indicating that on March 10, 1956, he had experienced a sudden onset of severe headache, dysphasia, and vomiting. He had no loss of consciousness, but lumbar puncture done that day showed bloody cerebrospinal fluid. His blood pressure was 260/180 and he developed a right hemiparesis shortly after the onset of symptoms. Until the time of admission he had shown gradual improvement with bed rest. Upon admission, examination indicated blood pressure of 200/120, marked hypertensive changes of the ocular fundi, right lower quadrantanopsia, minimal right facial paresis, mild right hemiparesis, and right hemihypalgesia. Skull films were negative, but electroencephalography indicated delta waves (Grade II) in the left temporo-parietal area. Genito-urinary examination, including cystoscopy, pyelograms, and aortography, suggested an uretral pelvic junction obstruction due to an anomalous vessel. At surgery the right kidney was found to be nearly destroyed and was removed. Following the nephrectomy on May 22, 1956, the patient's blood pressure returned to normal levels and his dysphasia showed striking improvement. The final diagnosis, confirmed in part at surgery, included intracerebral hemorrhage, hypertension, and uretral pelvic obstruction due to an anomalous vessel.

The psychologic testing was performed on April 30, 1956, shortly after the patient's admission to the hospital. In order to achieve complete objectivity in our procedure, our customary routine was followed. The patient was tested and the neuropsychologic report was written in the complete absence of any knowledge of anamnestic or clinical findings. We knew only the patient's name, age, education, handedness, and occupation. Examination for aphasia revealed pronounced dysphasia with symptoms including auditory verbal dysgnosia, dysnomia, dyslexia, and dysgraphia. The patient also had a mild tactile finger dysgnosia of the right hand. The receptive components of the dysphasia were particularly striking. The psychologic testing yielded poor performances on the verbal part of the Wechsler Scale, the Halstead Speech-sounds Perception Test, and Part *B* of the Trail Making Test. The patient had a verbal IQ of 88 but a performance IQ of 117. He made 27 errors on the Speech-sounds Perception Test whereas only 8 errors are sufficient according to Halstead's norms to place a subject's performance within the range characteristic of brain damage. He required 106 seconds to complete Part *B* of the Trial Making Test. The patient's results on these tests constituted his only poor performances and each of these tests explicitly require the use of language symbols. In many instances his other test results were remarkably good. His Category Test performance may be cited as an example. The patient made only 8 errors in responding to the 208 stimulus figures included in the test. This score is much better than the average achieved by normally functioning subjects without brain damage. This subject illustrates clearly that high-level performances requiring abstract thinking can be accomplished by persons with pronounced dysphasia. Although we cannot go into detail in this paper in explaining the interpretation of our battery of test results, the following "blind" neuropsychologic report was written just after the tests had been given.

> This 30-year-old man obtained a Full-Scale Wechsler-Bellevue (Form I) IQ of 102 (Verbal IQ, 88; Performance IQ, 117). The aphasia examination indicated the following symptoms: auditory verbal dysgnosia, dysnomia, dyslexia, dysgraphia, and mild right finger agnosia. This difficulty with the symbolic communicational aspects of language account for the disparity in verbal and performance IQ's. Although the subject obtained an Impairment Index of only 0.3, the W-B results

> and the dysphasic symptoms indicate definite brain damage which appears to be located maximally in the posterior part of the left temporal lobe and adjacent areas anteriorly. It would appear that our findings could be accounted for by postulating a vascular accident with rather focal effects and without much generalized involvement of the hemispheres. The Impairment Index of 0.3, together with excellent performances on certain tests, suggests that the immediate prognosis with respect to recovery of language functions is good. Personality measures indicate that this person is rather considerably disturbed affectively.

b. Patient E. W. This 45-year-old, left-handed, white male provides an instance in which we can compare the psychologic test results before and after the development of pronounced dysphasia. This patient first entered the hospital on April 17, 1958, with complaints of headaches on the left side, numbness of the right side of the body, and occasional episodes of difficulty with speech. The patient's symptoms had begun eight months prior to admission with tinnitus of the right ear, tingling of the right side of the body, dizziness, and blurring of vision. Early in the course of the illness the patient had experienced one episode of twitching of the right side of the face and right hand, lasting for 5 to 10 minutes. The patient had been unconscious for three days in 1942 as a result of head trauma in an automobile accident.

The neurologic examination indicated only a mild loss in auditory acuity of the right ear and a subjective impression of a possible loss in sensation on the left side of the body. X-rays of the skull, pneumoencephalography, and arteriography were reported as normal. Electroencephalography revealed delta waves (Grade I) of the left Sylvian area, but the tracings were felt to be only minimally abnormal. The patient was discharged on June 19, 1958, without a diagnosis but with instructions to return in one month for re-evaluation.

Psychologic testing was performed on May 7, 1958. Examination for aphasia revealed no positive findings. The report written at that time summarizes the results:

> This 45-year-old man obtained a Full-Scale Wechsler-Bellevue IQ of 119 (Verbal IQ, 114; Performance IQ, 122). The tests of biological intelligence yielded an Impairment Index of 0.6 which is consistent with mild dysfunction of the brain. An analysis of the details of the test results fails to give any definite indications with respect to focal involvement. Certain tests suggest some dysfunction in the left cerebral hemisphere and certain others imply dysfunction in the right cerebral hemisphere. The best conclusion would appear to be one of mild diffuse involvement. In spite of these findings on tests that are

> particularly sensitive to the organic condition of the brain, the relatively good performance on the Wechsler IQ Scale should be noted. The Minnesota Multiphasic Personality Inventory suggests that the patient has mild disturbances of an affective nature, but it would appear that these are not sufficient to be responsible for the relatively poor scores we obtained on tests sensitive to brain function.

The patient was re-admitted on August 20, 1958, with complaints of increasing headache, weakness on the right side, and visual difficulties. Neurologic examination revealed a mild right hemiparesis, papilledema, imperception on the right with bilateral simultaneous visual and tactile stimulation, hyperactivity of reflexes on the right, and pronounced dysphasia. Electroencephalography gave evidence of delta waves (Grade II) in the left temporo-parieto-frontal area. The previous films from angiography were reviewed and now interpreted as showing a definite tumor stain in the left Sylvian area. Craniotomy was performed on August 25, 1958, and a left temporal lobectomy performed. A tumor was found deep in the mid-temporal area. Histologic sections classified the tumor as an astrocytoma, Grade III.

The patient was re-examined psychologically on August 22, 1958, two days before surgery. This examination was also performed and interpreted in the complete absence of the anamnesis or clinical findings. The report of the neuropsychologic examination describes in some detail the changes that had occurred since the first testing.

> This 46-year-old man was tested originally on 5-7-58. At that time the test results were interpreted as indicating mild impairment of organic brain functions. He received a Full-Scale Wechsler-Bellevue (Form I) IQ of 119 (Verbal IQ, 114; Performance IQ, 122). The results presently obtained are very considerably poorer than those obtained initially. His present IQ on the same form of the Wechsler-Bellevue is 87 (Verbal IQ, 67; Performance IQ, 112). Because of limited available time not all of the tests of biological intelligence could be given, but the results obtained suggest that impairment of brain function is considerably worse at present than it was 3½ months ago. The patient's tapping speed in May was 46 with the right hand and 52 with the left hand. Presently the results indicate that he has slowed considerably with his right hand obtaining a score of only 37 although his tapping speed with the left hand was somewhat faster with a score of 58. His results on the Tactual Performance Test in May, 1958, were relatively poor but showed no particular lateralizing significance in comparing the performance with the two hands. At present, he was able to perform the task in 9.3 minutes with his left hand but was scarcely able to make any progress at all with his right

> hand in a period of 10 minutes. These results together with the results from the Wechsler-Bellevue Scale indicating marked impairment of verbal functions, are consistent in indicating serious structural damage in the left cerebral hemisphere. The aphasia examination indicated the presence of dysnomia, spelling apraxia, dysgraphia, dyscalculia, central dysarthria, right-left disorientation, visual number dysgnosia, visual letter dysgnosia, dyslexia, and auditory verbal dysgnosia. These symptoms clearly include ones of both an expressive and receptive nature, but the receptive losses were sufficiently pronounced to suggest a fairly posterior area of involvement in the left hemisphere. The patient also suppressed a tactile stimulus to the right hand when given simultaneously with one to the left hand, a stimulus to the right hand when given with one to the left face, and suppressed the stimulus to the right face when it was given in combination with one to the left hand. He also occasionally suppressed a stimulus to the right ear when an auditory stimulus was given bilaterally and simultaneously. Although the patient responded correctly to a visual stimulus to his right side or left side when given individually, he consistently suppressed the stimulus to the right side when given in combination and simultaneously with one to his left side. The patient also showed a finger dysgnosia for Fingers 2, 3, and 4 of the right hand although he identified fingers stimuluated on the left hand without failure. Finally, the patient showed a mild dystereognosis of the right hand but not the left. These latter results also suggest a fairly posterior location for the lesion. The level of impairment is sufficiently severe to suggest a destructive type of lesion that has fairly widespread significance in the hemisphere. Performances dependent upon the function of the right cerebral hemisphere were in general quite good. While the results are consistent with those we obtain from patients with intrinsic tumors or cerebral hemorrhages, the fact that functions dependent upon the right hemisphere are fairly good suggests that the likelihood of neoplastic disease is much greater than cerebral vascular disease. In conclusion, therefore, we would postulate that our results are quite consistent with an intrinsic neoplasm in the posterior part of the left cerebral hemisphere particularly involving the parietal and temporal regions.

The changes in findings on this patient, in comparison with his first examination, were quite pronounced. They involved rather particular areas, however, including tests requiring language functions and findings dependent upon lateralized motor and sensory function. The Wechsler performance IQ had dropped from 122 to 112, but the verbal IQ had changed from 114 to 67. Instead of being at approximately an average level or above on most of the verbal subtests of the Wechsler Scale, the patient's adaptive intelligence in verbal expression was below the limits of the Scale in many respects. This result is not surprising, of course, in consideration of the marked dysphasia noted in the neuropsychologic report quoted above. The patient had made 50 errors on the Category Test in the first examination and made 57 errors after his severe dysphasia had developed. He required 12.0 minutes to complete the Tactual Performance Test with his left hand on the first examination, but needed only 9.3 minutes on the second testing. Performance with his right hand was poorer on the second testing, however, in accordance with the development of the mild hemiparesis noted in the neurologic examination. In a number of other instances the patient's test performances were approximately at the same level as they had been initially, even though pronounced dysphasia had developed in the interval separating the two testings. The interesting point to be derived from comparison of the two psychologic examinations is that, excepting tests that explicitly require expressive and/or receptive language functions, the psychologic consequences of the rapidly developing tumor and dysphasia were not pronounced.

F. DISCUSSION

Direct inquiry of neurologists and neurosurgeons reveals that many have qualified opinions to the effect that brain lesions causing dysphasia have more seriously impairing effects on intelligence and thinking ability than do comparable lesions that do not cause dysphasia. This opinion is admittedly impressionistic in most instances. Except for the controlled experimental investigations referred to above, the fairly extensive literature concerned with this problem (see the reviews by Meyers 6, 5, and Reitan 8, 9) has also been based on clinical impressions. Even in the absence of satisfactory pertinent evidence, one may suggest that the opinions regarding the especially adverse effect on thinking and intelligence of dysphasia may very possibly reflect the fact that the clinical basis for estimating intelligence and thinking ability rests principally on the receptive and expressive aspects of verbal communication. Certainly, our results indicated that the dysphasic patients were especially impaired in this respect. They did no more poorly than the nondysphasic subjects, however, on even complex psychologic tests that did not explicitly require language functions.

The long-standing argument about the role of dysphasia with relation to intelligence measurements appears to receive a consistent answer from our results. Problem-solving and other adaptive abilities related to intelligence do not appear to suffer especial impairment in the dysphasic group as compared with the nondysphasic subjects, implying that language impairment is a relatively independent deficit rather than a direct manifestation of defective intellectual or thought functions.

As with most controlled experimental investigations, one must be cautious about generalizing beyond the characteristics of the samples compared. The groups were relatively small and we can have no assurance that they represented accurately the populations from which they were drawn. This realization is probably particularly important in evaluating the results with respect to the relation of language to thinking. The results suggest that language intactness is less important in complex thinking than has generally been presumed. We must point out, however, that we have not studied this question with patients having complete losses of language function. Our patients were clearly dysphasic, but each had some remaining facility in both receptive and expressive aspects of language function. In addition, the fact that dysphasic patients had in the past developed normal language functions may possibly have permitted implicit verbal reasoning at a higher level than may have been possible if they had never developed language functions. This consideration suggests that the problem should be investigated with groups of children who are as comparable as possible except that one group has failed to develop language functions while the other group has developed normally in this respect.

The fact that both brain-damaged groups generally performed more poorly than did the control group should also be considered in evaluating the relation of language functions to intelligence and thinking. The possibility exists that the nondysphasic group might have had unidentified but possibly even rather specific impairment different from that suffered by the dysphasic group which impeded thinking and intelligence to an extent comparable with the effects of dysphasia. Thus, dysphasia might have an impairing effect on thinking and intelligence even in tests not explicitly requiring language functions. The findings of this study suggest, however, that if this is the case, comparable brain damage without dysphasia produces equivalent impairment except on tests requiring language function.

G. Summary and Conclusions

The question of the relationship between language and thinking or intelligence was approached through evaluation of the performance of dysphasic patients on a wide variety of psychologic tests. Three groups of 32 subjects were used. One group was composed of patients with brain lesions and dysphasia, another of patients with brain lesions but without dysphasia, and the third of patients hospitalized for various reasons but with no anamnestic or clinical evidence of organic brain dysfunction. Subjects from the three groups were matched in triads with respect to color, sex, chronological age, and years of formal education in an attempt to equate the influence of these variables. In addition, patients in the two brain-damaged groups were matched fairly closely in pairs with respect to the type of brain lesion or disease that was present. Medical records and aphasia examinations for more than 1,500 patients were reviewed in order to compose the groups. An extensive battery of psychological tests was administered to each subject. These tests included the Halstead Neuropsychological Battery, the Trail Making Test, and the Wechsler-Bellevue Scale of Adult Intelligence (Form I).

The results indicated that both brain-damaged groups performed more poorly than the control subjects on all tests with the possible exception of the verbal subtests on the Wechsler Scale. On these tests the dysphasic group performed very poorly, but the nondysphasic group was not greatly inferior to the controls. Two additional tests specifically required use of language symbols and on both of these the dysphasic group performed significantly more poorly than did the nondysphasic group. On all additional tests, however, the dysphasic and nondysphasic groups had almost equivalent performances in spite of the fact that many of these tests were quite complex and would appear very possibly to be aided in their solution by implicit verbalization. The results suggest that language function is less important in complex thinking than has generally been presumed. Except for procedures specifically requiring the use of language symbols, adaptive abilities related to thinking and intelligence did not appear to suffer especial impairment in the dysphasic as compared to the nondysphasic group. This finding would appear to imply that language impairment in dysphasia is a relatively independent deficit rather than a direct manifestation of defective intelligence or thinking ability. Certain limitations of this study and conditions to be met in broadening the general significance of the findings were noted. However, the common impression that dysphasic patients are more severely and generally impaired intellectually than are patients having comparable brain damage but no dysphasia does not receive support from our results. Conversely, the results of this study should enhance the impetus for rehabilitational work with dysphasic patients in order to bring their language abilities up to the level of their performances on complex tasks not explicitly requiring language functions.

REFERENCES

1. DAVIDS, A., GOLDENBERG, L., & LAUFER, M. W. The relation of the Archimedes spiral aftereffect and the Trail Making Test to brain damage in children. *J. Consult. Psychol.*, 1957, **21**, 429-433.

2. HALSTEAD, W. C. Brain and Intelligence. Chicago: Univ. Chicago Press, 1947.
3. HEIMBURGER, R. F., & REITAN, R. M. Testing for Aphasia and Related Disorders. Indianapolis: Authors, 1957, (mimeo).
4. KIMBLE, G. A. Principles of General Psychology. New York: Ronald Press, 1956.
5. MEYERS, R. The relationship between "thinking" and language: An experimental approach using dysphasic patients. *Trans. Amer. Neurol. Assn.*, 1947, 65-69.
6. ———. Relation of "thinking" and language: An experimental approach, using dysphasic patients. *Arch. Neurol. & Psychiat.*, 1948, **60**, 119-139.
7. MUNN, N. L. Psychology (Second Ed.). New York: Houghton Mifflin Company, 1951.
8. REITAN, R. M. Intellectual functions in aphasic and non-aphasic brain injured subjects. *Neurology*, 1953, **3**, 202-212.
9. ———. Intelligence and language functions in dysphasic patients. *Dis. Nerv. System*, 1954, **15**, 1-8.
10. ———. The performance of aphasic, non-aphasic, and control subjects on the Rorschach Test. *J. Gen. Psychol.*, 1954, **51**, 199-212.
11. ———. Investigation of the validity of Halstead's measures of biological intelligence. *A. M. A. Arch. Neurol. & Psychiat.*, 1955, **73**, 28-35.
12. ———. The relation of the Trail Making Test to organic brain damage. *J. Consult. Psychol.*, 1955, **19**, 393-394.
13. ———. Validity of the Trail Making Test as an indicator of organic brain damage. *Percept. & Mot. Skills*, 1958, **8**, 271-276.
14. ———. The Effects of Brain Lesions on Adaptive Abilities in Human Beings. Indianapolis: Author, 1959 (mimeo).
15. REITAN, R. M., & TARSHES, E. L. Differential effects of lateralized brain lesions on the Trail Making Test. *J. Nerv. & Ment. Dis.*, 1959, **129**, 257-262.
16. ROSS, A. T., & REITAN, R. M. Intellectual and affective functions in multiple sclerosis. *A. M. A. Arch. Neurol. & Psychiat.*, 1955, **73**, 663-677.
17. WATSON, J. B. Behaviorism. New York: People's Institute, 1924.
18. WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1944.
19. WEISENBURG, T., & MCBRIDE, K. E. Aphasia: A Clinical and Psychological Study. New York: The Commonwealth Fund, 1935.

Department of Neurology
Indiana University Medical Center
1100 West Michigan Street
Indianapolis 7, Indiana

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 377-381.

A MULTIDISCIPLINARY APPROACH TO THE DIAGNOSIS AND THERAPY OF THE NONREHABILITATED ALCOHOLIC PERSONALITY*

Department of Education, Marquette University

CYRIL C. O'BRIEN

A. INTRODUCTION

The assumption of the existence of a nonrehabilitated alcoholic personality is predicated upon the clinical observations and statistical studies of individuals addicted to the abnormal consumption of alcoholic beverages. Such findings have been published by independent experimenters representing a diversity of disciplines (3, 5, 7, 20, 21, 23). Psychology, sociology, education, psychotherapy, economics, medicine, and nutrition are included among this varied list.

That there is a rehabilitated alcoholic personality has been indicated by data derived from the study of both overt and introspective behavior. The literature on the subject of alcoholism is replete with examples of reinstated alcoholics (1, 2, 8, 13, 15, 16).

Although the word *personality* has been defined in many ways and is regarded by some as an ambiguous term, the connotation assigned to it here comprises not only the sum total of all traits that differentiate one individual from another, but also the interaction, coordination, and integration of such characteristics within the individual with a special consideration of the factors of volition, conation, cognition, and affectivity. The acceptance of this definition implies the recognition of the complexity of the *persona*. Although the classifying of a particular personality as alcoholic delimits the field, it does not necessarily make those phases of personality less complex (12, 17, 18).

While the dichotomous classification of rehabilitated and nonrehabilitated alcoholic may elicit few objections, if any, the typing of *alcoholic personality* may well produce a demand for verification. One may taxonomize a nonrehabilitated alcoholic personality on the grounds that individuals so typed frequently resort to rationalizations, indefensible alibis, and escape mechanisms of a puerile and infantile kind in order to justify their drinking.

---

* Received in the Editorial Office on August 30, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The nonalcoholic personality is differentiated in that while it may show similar childish evasions and attitudes, it does not reveal indications of compulsive drinking (5, 10, 14, 23, 25). However, in the case of a lifetime abstainer or teetotaller, there is no known way at present of accurately predicting if drinking would initiate a compulsion. The evidence that a system of rationalizations alone will set up the compulsive drinking pattern experienced by the alcoholic is inconclusive (2). Still the compulsive drinking and rationalizations require both a physiological and psychological diagnosis (2, 6, 20, 22, 24).

B. Types of Drinking

The alcoholic personality is so constituted that the imbibing of alcohol in any form incites a physiologically compulsive drinking pattern (1, 2, 6, 10, 25). There is also evidence that some alcoholic personalities have, in addition, a psychological compulsion (2, 13). Such interior drives may be increased by social factors affecting the drinker. Prominent among these factors are unfortunate living conditions.

Nevertheless, to escape undesirable living conditions is only one reason why people drink. Just to be sociable is often cited as the main motive. Explanations for imbibing include a release from daily restrictions, an escape from frustrations, and an antidote for worry (8, 19). Psychologists and the followers of other disciplines are interested in the scientific study of alcohol as a social lubricant. They are also concerned when drinking becomes excessive or compulsive. Abnormal drinking can ramify from effect on family and friends to involvement of community, police, and courts (3, 7, 10, 11).

Not only legal but moral concomitants are usually subsumed under an alcoholic personality (1, 8, 9, 17). In a plan for rehabilitating the active alcoholic, the Twelve Steps of Alcoholics Anonymous emphasize the moral. In Step Number 4, alcoholics are admonished to "make searching and fearless moral inventory of themselves (1)." A self-examination helps the alcoholic realize the nature of compulsive drinking (1).

Considering the problem conatively within a philosophical frame of reference, an individual might be impelled without being compelled to drink. By virtue of being a social or moderate drinker, one is not compelled to drink, but on occasion may be impelled to imbibe, yet knowing when and being able to stop (12). Re-education consists in changing attitudes and in showing the alcoholic personality how not to begin to drink (1, 5, 8, 14).

C. SUGGESTED HYPOTHESIS

The hypothesis is presented here that the compulsive drinking pattern of the alcoholic has its etiology in a combination of biochemical, endocrinological insufficiencies. These conditions force the addicted individual to seek defense reactions that help consolidate undesirable attitudes. As an initial theoretical approach to the understanding of the alcoholic personality, the alcoholic's affinity for alcohol could be explained as being partly due to the asynchronization of the endocrine glands caused by a deficiency of diffusible calcium in the blood stream resulting in lower calcium metabolism and lack of replenishing calcium to the anatomical structure and system of the individual as a whole. An analogy might be drawn to the recent discovery of the relationship of blood chemistry to schizophrenia. The alcoholic shakes that frequently follow a compulsive drinking bout can generally be eliminated by the intravenous injection of intensive calcium therapy (13, 15, 16). Research indicates that the passage of the nerve impulse from nerve to muscle across the myoneural junction is facilitated by the remedying of calcium deficiency (2, 4).

Other conditions of the nonrehabilitated alcoholic personality associated with a physiological and biochemical origin comprise neuro-muscular irritability, blood chloride imbalance, hypoglycemia, increased permeability of cell structures, lactic acid excesses, and anoxia (2, 21, 22). The rationale for the therapeutic control of the foregoing is largely pharmacological in nature (2, 21, 22).

D. A MULTIDISCIPLINARY APPROACH

Several studies have established that alcoholics require medical attention (6, 11, 13, 15, 16). Many compulsive drinkers can be referred directly to the counseling therapy of Alcoholics Anonymous and as a sole result of such become rehabilitated (1) or receive the dual benefits of psychological diagnosis and psychotherapeutic aid culminating in rehabilitation (1, 8, 10). More often than not, however, the nonrehabilitated alcoholic personality requires the services of those of varied prefessional backgrounds. Such a personality encompasses the physical, the physiological, the neurological, the biochemical, the endocrinological, the psychological, the social, the moral, and the educational aspects (2, 15, 16, 17).

At least in the more difficult cases of disturbances within the personality structure of the alcoholic, the ministrations of several of the following are necessary: the nurse, the physician, the psychologist, the psychiatrist, the pastor, the social worker, the nutritionist, and the educator. This multidisciplinary approach tends to harmonize with the concept that alcoholism is a disease embracing physical, mental, social, and spiritual areas of activity (8, 14, 19).

With the exceptions of the more or less constant qualitative elements of physiological compulsion and frequent rationalization, the nonrehabilitated alcoholic personality is highly variable in the other traits and components comprising the personality in toto. The rehabilitated alcoholic personality is also strongly deviant in such components, but through the management of its alcoholic addiction, has removed—so long as he refrains from alcoholic imbibing—the physiological compulsiveness to continue drinking.

The psychologist as an indispensible member of a multidisciplinary team, as a diagnostician and psychotherapist, can use scientific skills in determining whether a specific case of a nonrehabilitated alcoholic personality is an instance of alcoholism *per se,* or whether there are in addition to the alcoholic compulsions and rationalizations other deviations of personality that warrant therapy.

E. SUMMARY

This paper is concerned with (*a*) a rationale for the classification of an alcoholic personality, (*b*) the differentiation of a nonrehabilitated from a rehabilitated alcoholic personality, (*c*) the significance of a physiologically compulsive drinking pattern, (*d*) constant psychological characteristics and attitudinal responses associated with the alcoholic personality, and (*e*) certain social, moral, and philosophical implications of the problem.

An initial theoretical approach to the understanding of the alcoholic personality is presented in terms of an hypothesis embodying biochemical, endocrinological, and physiological concomitants. A reasoned exposition outlines the need for specialists from a diversity of disciplines including psychology and sociology to deal with the alcoholic personality.

REFERENCES

1. Alcoholics Anonymous. 44 questions and answers about the program of recovery from alcoholism. New York: The Alcoholic Foundation, 1952.
2. ARNOLD, W. E. Anoxic neuropathic diathesis. Detroit, Michigan: Atlas Pharmaceutical Laboratories, 1951.
3. BACON, S. E. Sociology and the problems of alcohol. Memoirs of the Section of Studies on Alcohol, Yale Univ., 1944.
4. CANTAROW, A. Calcium metabolism and calcium therapy. Philadelphia: Lea & Febriger, 1933.
5. HIRSH, J. Alcohol education. New York: Schuman, 1952.

6. JELLINEK, E. M. Phases in the drinking history of alcoholics. Memoirs of the Section of Studies on Alcohol, Yale Univ., 1946.

7. LANDIS, B. Y. Some economic aspects of alcohol problems. Memoirs of the Section on Alcohol Studies. Yale Univ., 1945, No. 4.

8. MIELCAREK, H. A., & O'BRIEN, C. C. The problem drinker in industry. Milwaukee, Wis.: Allis-Chalmers Mfg. Co., 1950.

9. O'BRIEN, C. C. Alcoholism as an industrial problem. *Midwest Sociol. Rev.*, 1949, **12**, 14.

10. ———. Alcoholism among disciplinary cases in industry. *Quar. J. Stud. Alcohol,* 1949, **10**, 268-278.

11. ———. Alcoholism among disciplinary cases in industry. *Digest Neurol. & Psychiat.*, 1949, **17**, 357.

12. ———. Excessive drinking versus industry. *America,* 1949, **80**, 654-655.

13. ———. Calcium therapy in the treatment of alcoholism. *Amer. J. Pharmacy,* 1951, **123**, 154-159.

14. ———. Education on alcoholism. *Amer. Cath. Sociol. Rev.*, 1952, **13**, 107-110.

15. ———. Experimental evidence in the treatment of alcoholism by intensive calcium therapy. *J. Amer. Osteopath. Assoc.*, 1952, **51**, 393-394, and 426.

16. ———. Hospital and ambulatory cases of alcoholism: intensive calcium therapy. *J. Amer. Inst. Homeopathy,* 1952, **45**, 83-84.

17. O'BRIEN, C. C., & STRACHAN, G. J. (*Ed.*) Proceedings of the institute on alcoholism. Milwaukee, Wis.: Marquette University, 1953.

18. O'BRIEN, C. C. Designing a college course in alcoholism. *Amer. Cath. Sociol. Rev.* 1954, **15**, 265-267.

19. SELINGER, R. V. Alcoholics are sick people. Baltimore: Alcoholism Publications, 1945.

20. WAELDER, R. Basic Theory of Psycholanalysis. New York: International Univ. Press, 1960.

21. WILLIAMS, R. J., BERRY, L. J., & BEERSTECHER, E., JR. Individual metabolic patterns, alcoholism, and genetotropic diseases. *Proc. Nat. Acad. Sci.*, 1949, **35**, 265-271.

22. WILSON, A., & SCHILD, H. O. Applied Pharmacology. Boston: Little, Brown. 1959.

23. WILSON, H. H. Management of alcoholism. *California & Western Med.*, 1936, **45**, 349-352.

24. WOODS, R. Nutrition and alcoholism. *Borden's Rev. Nutrit. Res.*, 1952, **18**, 1-48.

25. WORTIS, H., SILLMAN, L. R., & HALPERN, F. Studies of Compulsive Drinkers. New Haven: Hillhouse Press, 1946.

2568 North Summit Avenue
Milwaukee 11, Wisconsin

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 383-397.

MARRIAGE AND PREMARITAL CONCEPTION*

Mental Hygiene Clinic, Santa Barbara, California

HORACE GRAY

A. AIM

Marriage among adolescents is a grave psychological problem to them and a social problem to their parents and the community. On the one hand the young are chided for petting proclivities, while on the other hand adults commonly forget the behavior of their own contemporaries in youth. The flood of books and articles in family magazines are conspicuously evasive or ignorant. A first aim therefore must be to discover and uncover current psychological attitudes and courtship behavior. Then the problems may be faced and worked on more fruitfully.

B. SOURCES OF DATA

These are partly brief references to important findings in print and mainly tabulations from marriage and birth certificates in the office of the Recorder of this county (38). It has a population of something over 100,000, about half urban and half rural cattle ranches. Marriage certificates are filed at the county seat where the license was taken out, regardless of the place of marriage. These filings therefore include persons licensed here but married in another county, and lose those licensed elsewhere but married here. The numbers in these two categories, and the reasons, might be instructive, but are not readily ascertainable, and may be assumed to balance, without significant effect on this study. The marriage certificates were turned through consecutively, filed in the order in which received from the offices of the city health officer (36) and the county health officer (37). When the bride was under her 21st birthday, the data were copied.

C. PERIOD SURVEYED, AGE, GRADE, AND MARITAL STATUS

1. *Period Surveyed*

This was chosen as from June 1, 1956, through July 31, 1958, i. e., 26 months or more exactly 789 days.

---

* Received in the Editorial Office of the Journal of Social Psychology on February 1, 1960 but eventually transferred to this Journal. Copyright by The Journal Press.

2. *Age of Bride*

Attention is focussed on the brides, since the majority of the grooms were over 21. Since age was recorded as latest birthday, age 14 means 14 but less than 15, midpoint 14.5 years, and so forth. The distribution by age is displayed in Table 1. There were 907 such marriages, which is 39.9 per cent of the 2274 marriages of all ages during the period. The mode (most frequent value) was the age-group centered at 18.5, and the arithmetic mean was 18.7 years. It may be noted that the numbers at 19 and 20 are lower than at 18. As the reason for the concentration at age 18, one may speculate: (*a*) that graduation from high school is a natural turning-point to commencement of a new way of life, or (*b*) the legal age for marriage in a majority of states leads to over-statement of age (Christensen, Andrews, and Freiser, 19). In this state it is required that a girl under her 18th birthday must have approval of a parent, and under 16 must have the assent of a superior court judge as well.

At the lowest ages the following quotations are apposite: "When Jesus was born, Mary was probably fifteen years old" (Bamm, 3). On this estimate she was only about 14 at conception. Again, "She hath not seen the change of fourteen years, a fortnight short of 14; let two more summers wither in their pride ere we may think her ripe to be a bride"; i. e., after her 16th birthday (Shakespeare, 41). It has been ironically remarked that this heroine whose role so many actresses have yearned to play, would by modern standards be regarded as a juvenile delinquent.

3. *Age of Bride, Comparisons*

From the table the most striking features among married girls under 21 years of age are obtained by comparing Rows II, VII, XI, XIII. It is seen that the excess married at each age under 18, versus 18 and over, is greater in the Santa Barbara and Dallas (White, 51) samples than in the estimates based on the recent census (Brunsman, 48).

Comparing Row III with IX, last column, for all ages, a larger percentage were married under 21 in the Iowa sample (Monahan, 33) than in the Santa Barbara sample.

Another comparison between the same two rows, II and IX, where those married at each stated age, is shown in percentage of marriages at all ages. By adding across the proportions for marriages under 18th birthday pooled, our sample shows a higher frequency, 11.7 per cent, versus 8.6 per cent in the Iowa sample.

TABLE 1
DISTRIBUTION OF FEMALES BY AGE AT MARRIAGE (LATEST BIRTHDAY)

	14	15	16	17	18	19	20	Total under 21 y.	
This sample									
No. by age	5	13	89	159	254	188	199	907	I
% of 907 married at each age	0.6	1.4	9.8	17.5	28.0	20.7	21.9	100	II
% of 2274 married all ages, at each	0.2	0.6	3.9	7.0	11.2	8.3	8.7	39.9	III
S.B. 1950 Census, Brunsman									
No. by age	234	235	243	254	371	401	356	2094	IV
% of 2094 pop. at each age	11.2	11.2	11.6	12.1	17.7	19.2	17.0	100	V
Ever-married, no. each age	1	4	14	36	101	168	192	516	VI
% of 516 ever-married, by ages	0.2	0.8	2.7	7.0	19.6	32.6	37.2	100	VII
% of females at each age to be ever-married	0.4	1.7	5.8	14.2	27.2	41.9	53.9	24.6	VIII
Iowa, Monahan '53									
% of wives all ages, married at stated ages	0.1	0.6	2.3	5.6	14.5	14.2	13.5	50.8	IX
Dallas, White '59									
Married girls of each age	12	57	103	166	112	18	8	476	X
% of the 476	2.5	12.0	21.6	34.9	23.5	3.8	1.7	100	XI
Nebraska, Cannon '59									
No. girls at each grade (age)	1381	1099	293	683				3456	XII
% of 3456	40.0	31.8	8.4	19.8				100	XIII
% of 3456 married in each grade	0.4	1.0	0.6	3.0					XIV
% of each grade who married	1.0	3.1	7.2	14.9					XV

This pool under 18th birthday amounted to 29.3 per cent of the 907 marriages under 21st birthday, i. e., the dangerously youthful unions were much more frequent than the 10.7 per cent obtained from the census estimate in Row VII (Brunsman, 48).

4. *Grade*

In passing it may be noted that grade is a less precise grouping than age, since the age-composition of any grade is either unknown or at most is variable between high schools reported. For those who think more readily in terms of grade, the following brief statement may be worth reciting: Among 60 Iowa girls married before graduation from high school, 70 per cent were in the 12th grade, 22 per cent in the 11th, and 8 per cent in the 10th. These marriages were disapproved by 25 per cent of the brides' mothers and by 23 per cent of their fathers (Burchinal, 6).

5. *Marital Status*

At the time of the marriages of record, 98.3 per cent of the brides were single, none widowed, 1.0 per cent divorced, 0.7 per cent annulled. The latter two categories, amounting to 1.7 per cent of the 907, are here not further pursued, because divorce proceedings show mainly the vague and monotonous complaint as mental cruelty and because annulments are not open to study. Although the proportion of marriages at each successive age to end in divorce cannot be adequately documented from our sample, it is very important to remember that the rate is vastly highest under the 15th birthday, decreasing at 15, 16, and 17, but not approaching the adult rate until the bride has passed her 18th birthday (Monahan, 33).

Comparison may be made with census data on 236,480 females under 21 years of age: 94.4 per cent single, 0.1 per cent widowed, 2.3 per cent divorced, 3.1 per cent not stated. The notable difference is that fewer of our sample had been through a prior marriage (U.S. Vital Stat., 1956, 45).

6. *Groom's Age*

He was older most commonly by two years, averaging 3.51 years, with a standard deviation (for what it might be worth in this skewed distribution) of 2.86 years, and hence with standard error of the mean of about 0.1 year. Thus our grooms average slightly older than the figure of 3.3 years given for women under 30 (Glick, 22).

The husbands were one or two years younger than their brides in 28 instances or 3.1 per cent of the 907. Such situations are supposed to be unfavorable, but evidence on this point is not available from our sample for lack of information on the girl's motives for marriage and on the success or failure of the marriages.

When we consider the age spread in relation to the successive ages of brides, we find that the grooms averaged 6.0 years older than brides of 14, 4.2 years older than brides 15 and 16, 3.5 years older than brides of 17 and 18, and finally, close to 3 years older than brides of 19 and 20. Clearly, the younger the bride, the greater the tendency to marry a fellow a good bit older than herself, as has indeed been previously reported (Burchinal and Chancellor, 4).

When we consider the girls pregnant at marriage, the mother tended to be older than the father. Thus, in a sample in which the interval from marriage to birth was less than 7 lunar months (196 days), the increasing tendency could be shown by this arrangement: When ages same, 14 premarital conceptions out of 136 unions or 10.3 per cent; when wife younger 140/1207 or 11.6 per cent; and when wife older 24/188 or 12.8 per cent (Christensen, 17).

D. Interval from License to Marriage

The marriage-date followed the license-date by a wide variety of intervals, from 0 (same day) to 254 days. The frequency-distribution is shown in Table 2, expressed in percentage of all 907 marriages. It is surprising that so large a fraction of the couples, 78 per cent of 907, applied for license only within a week of marriage. The commonest interval (mode) was 0 (same day), the average was 5.4 days. By ages, the youngest were in the most hurry, as shown by the mean intervals for the younger ages, 14 through 17, in Table 3.

TABLE 2
Interval from License to Marriage Percentage of 907 at Each Interval

0	1	2	3	4	5	6	Later	Total
22.5	13.2	12.1	11.2	7.4	8.2	3.5	21.9	100

TABLE 3
Interval from License to Marriage Average Number of Days for Brides of Various Ages

Age	14	15	16	17	18	19	20
Days	0	4.5	2.8	3.8	6.2	5.2	7.2

E. Interval from Marriage to Birth

Now we come to the core of this paper, the interval to first birth of record, which may be noted as MB1. The words "first of record" probably do not tell us truly the first birth to many of the couples, as will be discussed later. Also, at a later time, our data may be worth examining with regard to the relationship of MB1 to the duration of pregnancy. The latter item can be computed from conception to first birth, on the common assumption that the closest estimate of date of conception is the first day of the latest menstrual period, which is ordinarily stated by the mother on the birth certificate. The reliability of this statement is questioned by some obstetricians, but despite reservations, may prove worth presentation.

The distribution of MB1 was made from counts of the IBM cards in class-intervals of 14 days, then for greater conciseness, compressed into 28 classes and shown in Table 4. The range runs from 364 days, or 1 year *before* marriage, to 999 days or 2.74 years after marriage.

If we plot the frequencies (vertical) against the class-intervals in months, from the greatest minus at left, to the greatest plus value at the right, we find a roughly bell-shaped curve with mode at 209d, median at 267d, and mean at 315d. This distribution is worth putting on record. The really important items are these:

(*a*) The births *before* marriage: 2 per cent of the N294 in S.B., and none in Iowa series of N660.

(*b*) The births before seven lunar months (196 days): 24 per cent and 27 per cent respectively; good estimate of definite premarital conceptions in our culture. This is the low limit of the true incidence.

(*c*) The births before eight lunar months (224 days): 43 per cent and 38 per cent respectively. These are less conservative limits, and undoubtedly a nearer approximation to the truth.

(*d*) The births before nine lunar months (9 × 28 days, i. e., less than 252 days: 55 per cent and 49 per cent respectively, average say 52 per cent as a probably still nearer approximation.

(*e*) The births before 10 lunar months (280d): 65 per cent and 58 per cent respectively, estimates of the high limit of possibly premarital conceptions.

(f) The births within one year of marriage: 35 per cent and 43 per cent, both much lower than the ordinary expectation of a first baby, and therefore suggesting that the first-births-of-record fail to reveal earlier first births not found in the records of the county of marriage. How many are these unknown true first births needs further investigation.

TABLE 4
INTERVAL MARRIAGE TO BIRTH (Santa Barbara and Christensen, 17)

Interval MB1 (minus means before marriage)		Mothers under 21			All ages
in days (1)	in 28-d. months (2)	Santa Barbara f (3)	Santa Barbara % of 294 (4)	Iowa % of 660 (5)	Iowa % of 1531 intervals, incl. over 1 year (6)
—364 to —337	—13	1	0.3		
—336 —309	—12				
—308 —281	—11				
—280 —253	—10	1	0.3		
—252 —225	— 9				
—224 —197	— 8				
—196 —169	— 7	1	0.3		
—168 —141	— 6				
—140 —113	— 5				
—112 — 85	— 4				
— 84 — 57	— 3	1	0.3		
— 56 — 29	— 2				
— 28 — 1	— 1	2	0.7		
Sum before marriage		6	2.0	0	0
0 to 27	1	0	0	2.0	
28 55	2	2	0.7	2.0	
56 83	3	6	2.0	1.2	
84 111	4	7	2.4	2.6	
112 139	5	9	3.1	3.0	
140 167	6	16	5.4	6.4	
168 195	7	25	8.5	9.8	
Sum under 196 d.ff.M.		65	22.1	27.0	
Cumulated-sum, incl. births before marriage		71	24.2	27.0	11.6
196 223	8	55	18.7	11.2	
Cum. sum under 224 d.		126	42.9	38.2	16.5
224 251	9	37	12.6	10.6	
Cum. sum. under 252 d.		163	55.5	48.8	21.0
252 280	10	29	9.9	9.7	
Cum. sum. under 280 d.		192	65.3	58.5	25.2
280 307	11	38	12.9	15.3	6.6
308 335	12	35	11.9	13.9	6.1
336 363	13	29	9.9	12.3	5.3
Sum 280-363 days		102	34.7	41.5	18.0
Cum. sum. under 1 yr.		294	100.0	100.0	43.1

The foregoing figures may be interpreted as evidence that premarital conceptions in youthful brides were somewhat more common in this Santa Barbara sample than in the Iowa sample. The difference, whether the cut-off be taken at seven completed lunar months (*a* plus *b*), or eight months (*c*) or nine 9 months (*d*), or 10 completed lunar months (*e*), was about 12 per cent. The higher frequency in our sample may be due to either the approximately 5 years' later period, or to the local culture, or both.

A second comparison shows the frequency percentage of the interval from marriage to birth, MB, to be expected in percentage of births to mothers of *all* ages, yields still more striking evidence of the excess of births in young mothers. We did not tabulate births to mothers of all ages, but Christensen did (17), and the figures for all ages in Col. 6 may be compared with his and our figures for teenagers in Col. 4 and Col. 5. For all ages, we follow same arrangement as above for the adolescents:

(*a*) Before marriage are not given by him.

(*b*) Before seven months: 11.6 per cent of his all age N1531, i. e., markedly lower than the teenage percentages—around 22-27 per cent.

(*c*) Before eight lunar months: 16.5 per cent of his all age N1531, i. e., markedly lower, by about 41 per cent than the teenage percentages around 43-38 per cent. For all ages, Croog, '52 shows figures for Denmark, Finland, and Sweden which, if I understand aright, run about 29 per cent, under 8 months.

(*d*) Before nine lunar months: 21 per cent of his all age N1531, i. e., again about 40 per cent of the teenage percentages which were 55 and 49 per cent.

(*e*) Within 10 lunar months: 25.2 per cent of all ages, i. e., markedly lower than teenage percentages of 65-58 per cent.

(*f*) Within 1 year of marriage: 43.1 per cent of all ages, which, as expected, is *not* lower than the teenage percentages of 35 per cent and 42 per cent.

A third comparison may be of interest. From a family genealogy, taking the 40 first births between 1708 and 1943, we find a distribution of MB, ranging from 259 to 719 days, with mode 370 days, mean 375 days, and standard deviation 9.9 days. If we assume that conception occurs on the marriage night, and take values under 280 days, we find 247 days, 273, 276, and 278; these are short of 280 by 33, 7, 4, and 1 day, which in terms of deviation divided by SD, are found from a probability table likely to occur in 0.04 per cent, 24 per cent, 34 per cent and 46 per cent of all cases. It may be concluded, as one would expect, that in this earlier period, premarital conceptions were much less frequent than today.

A further approximation to the actual frequency of premarital conceptions may be offered. Let us consider, not alone the 26 definite cases mentioned already, but scrutinize the frequencies of MB1. The average interval from marriage to first birth certificate, MB1, for all the 414 mother-child pairs in the sampling, has been shown as 315 days. And let us assume that any MB notably greater than that average suggests the presumption that the woman bore a child—but unreported in this county—before the first child in our birth list, called B1. In other words, let an MB1 of 315 days be a presumptive cut-off point. As the nearest class-interval high-limit of 307 days, there are 184 values beyond that, or 44 per cent of the 414; these may be tentatively presumed as an approximation to the quantity of lost first births.

F. LOST OR CONCEALED FIRST BIRTHS

The six cases or 1.45 per cent of the 414 of record is so low that many of the longer intervals fail to record how many of such first-births-of-record really are second births. Causes to be considered are: (*a*) low fertility of the couple; (*b*) pregnancies ending in a still-birth, recorded but not picked up in this study; (*c*) first births reported in some other counties, and therefore not discovered in this study; (*d*) miscarriages from disease, such as nephritis or syphilis; (*e*) miscarriages due to accident; (*f*) legal therapeutic terminations; or (*g*) illegally induced abortions.

As to (*a*), relative infertility is fairly common, though I have found no figures. Not rarely does a couple after several years of unsuccess adopt a child, after which the wife bears two or more children. Perhaps relevant here are census data on many thousands of women giving birth for the first time, which showed that 32.5 per cent did so in the year following marriage and 32.8 per cent in the second year (Schachter and Grabill, 39). And, when region is considered, an amazing difference has been detected: among Cornell alumnae 20 per cent within the first year versus 58 per cent among Mormons (Christensen and Bowden, 18). Complete sterility occurs in about 12 per cent (Meaker, 32) to 17 per cent (Haman, 23) of marriages.

As to (*b*), stillbirths, we are not now concerned.

As to the remaining items, the proportion of pregnancies failing to end in live birth was found in a survey in Chicago, on white women, to be 11 per cent. This was closely matched by 10 per cent from Indianapolis, divided thus: unintentional 7.5 per cent, stillbirth 1 per cent, therapeutic 0.5 per cent, illegal 1 per cent. The same authority stated that in 1947 over 40 per cent of illegitimate births were to girls under 20 years of age (Dublin, 21). This leads us to inquire further into the relationship of illegitimate live births to age of mother. The published ratios (U.S. Vital Stat., 44, 45, 46) per thousand live births have been converted by moving the decimal point to show the ratio in per cent of live births in each age-group, which is a manner more familiar to most readers, in Table 5. We observe the extraordinary high value of 66 per cent for mothers under age 15, 23 per cent for mothers of 15 and less than 18 years, and 10 per cent for mothers aged 18 and 19. At older ages, with which we are not concerned, the figure is much lower, averaging around 3 per cent.

TABLE 5
ILLEGITIMATE LIVE BIRTHS, RATIO TO TOTAL LIVE BIRTHS
(Converted from per thousand as published in U.S. Vital Statistics (44, 45, 46) to percentages as being more convenient for most readers. Three years given to show the constancy of the striking high values under age 20.)

Age of mother	1955	1956	1957
Under 15th birthday	66.3	66.1	66.1
15 through 17	23.2	23.0	23.1
18 through 19	10.2	10.0	9.8
20-24	4.4	4.4	4.4
25-29	2.5	2.6	2.6
30-34	2.2	2.3	2.5
35-39	2.4	2.5	2.6
40 and over	2.6	2.6	2.9
All ages	4.5	4.6	4.7

Finally, as to (*g*), illegal abortions, there are the careful tabulations by a specialist (Timanus, 42) who, with one other abortionist, shared 90 per cent of the Baltimore practice. The mothers under 21 were 705 or 13 per cent of this series of 5,210 illegal operations, almost five times as many as turned up in the third Kinsey report (28). Furthermore, among these 705 cases were 28 or 4 per cent who had had a previous abortion. It was also estimated that one man did about 1,000 abortions a year, over a period of about 50 years. Among all the women in the sample ever married, 10 per cent had had induced abortions by 20 years of age.

G. UNPLANNED PREGNANCIES AND AGE

The proportion of unplanned pregnancies was greater among wives who conceived early in marriage. In a sample of 212 wives in Michigan State College in 1949, 36 per cent had planned conception, 26 per cent neither tried to achieve nor to avoid conception, while 38 per cent had tried to avoid conception (Poffenberger, 35). Evidently their information on birth-control was inadequate, or their efforts not very diligent; the best method, diaphragm and jelly, should yield less than 5 per cent failures.

H. PREMARITAL CONCEPTIONS AND AGE OF MOTHER

The number of cases shown in the total for Column 3 of Table 4 was 294, because there the purpose was to display the distribution of births within a year after marriage. For the present purpose, however, Table 6 shows a total of 414 first births to brides in our sample period, namely from June 1, 1956, through July 31, 1958. The additional 120 births were at intervals above the first year, as picked up in listing all first births to these mothers, discovered in listing births from December 1, 1956, i. e., six months after the earliest marriage through July 31, 1959, i. e., one year after the end of the marriage survey period. Therefore, a certain number of girls married in the earlier part of the said period had the first child more than one year after marriage, possibly as long as three years and two months after marriage. Such instances account for the additional 120, or for the total of 414.

TABLE 6
PREMARITAL CONCEPTIONS AND AGE OF MOTHER

Age (last)	14	15	16	17	18	19	20	Total
Total N	4	12	51	87	113	85	62	414
% of 414	1	3	12	21	27	21	15	100
Under 7 mos. N	2	9	15	14	15	12	4	71
% of 71	3	13	21	20	21	17	6	100

When the percentages in the fifth row are compared with those in the third row, it becomes evident that the frequency of premarital conceptions is markedly greater among girls who marry before their seventeenth birthday.

I. BIRTHS BEFORE MARRIAGE

There were six instances of pre-marital births. Range 364 to 13 days before marriage, mean 152 days. Four were 18 years old and two were 19. The number of definite premarital conceptions may perhaps be enlarged by counting mothers who stated on the marriage license that they were single, apparently falsely, because on certificate of first birth dated after marriage stated that they had had a prior living child. There were 26 such mothers, and 11 of these set the date of latest menstrual period earlier than that of marriage. On this evidence, conception occurred before marriage by an average of 129 days, ranging from 7 to 340 days. The ages of the 11 mothers were as follows: 1 at 14, 1 at 15, 3 at 16, 1 at 17, 3 at 18, 1 at 19, 1 at 20, with an average of 18.3 years.

J. PREMARITAL COITUS

In the literature may be found relevant data, though scattered, and not in uniform terms. Distinction was not always clear between intercourse with fiance and intercourse with previous suitors. The following are roughly in order of date of publication, not in order of magnitude of percentage frequency.

The statistics put out by the Registrar-general for England and Wales for the year 1939 showed that of the 20,981 women who gave birth to the first legitimate child, 15,427 or 73.5 per cent became mothers within the first eight months of marriage, and 14,787 or 70.5 per cent within the first seven months. If one adds the 5,219 illegitimate births to mothers under 20 during the same period, one finds that about 80 per cent of these women were unwed at time of first sexual intercourse (Landau, 29).

Among 296 girls married before the 19th birthday, 110 or 37 per cent were reached for interview by a married woman, and of these, 31 per cent admitted to pregnancy at time of marriage (Moses and Gingles, 34).

Out of 615 couples 274 or 45 per cent had intercourse with spouse before marriage (Burgess and Wallin, 13). A very similar figure, 44 per cent, was found in a study of 177 wives of students on campus, the wives averaging 2.5 years of college themselves (Kanin and Howard, 27). These observers made a further statement which supports the Kinsey studies (28) as to the influence of social stratum, namely, that when the student-wives were divided according to Warner's (50) well-known scale, and because of the smallness of the sample pooling the lower two classes, the premarital experience of the women rose to 83 per cent.

In an elaborate investigation in New Haven, their class V, which contained like Warner's scale about 23 per cent of the households, contained 40 per cent pregnant at marriage (Hollingshead and Redlich, 25).

Of 1425 marriages in high schools in California, at least 44 per cent were forced by pregnancy (Landis, 31).

Of 58 Iowa high-school brides, 40 per cent were premaritally pregnant (Burchinal, 8).

In Copenhagen in 1958, for those who became mothers under age 20, 90 per cent of first births took place within 9 months of marriage and 97 per cent within the first year. The intervals were grouped somewhat differently from ours, but are usable: 3.4 per cent within the first month of marriage, 6.0 per cent during the second month, 71.5 per cent during the third through sixth months, 12.5 per cent during the seventh through ninth months, and 3.4 per cent during the eleventh and twelfth months (Holm, 26).

The Youth Council of the Commissioners of Washington, D.C., are quoted as finding that in the seven months from September, 1957, through April, 1958, there were 185 pregnancies among unwed girls in the schools of the capitol (Am. Med. Assn., 2).

A TV documentary film (43) may attract public attention to little-known but apparently rather carefully searched-out material. A study of lonely-hearts clubs and marriage agencies was made by NBC's Eve Sandeffer, who in the guise of a widow enrolled in several marriage clubs. Scientific match-making has been seriously studied for several years, and indeed operated as a Personal Acquaintance Service, without profit, by the chairman of the sociology department at Los Angeles State College (Wallace and Odell, 49). Returning to the documentary, that cited a protestant episcopal minister to the effect that 75 per cent of brides were pregnant; and that 90 per cent of applicants for licenses and the ceremony were so eager as to claim that they were pregnant, with no evidence or possibly with only a statement from some venal physician.

K. SUMMARY

A study was made of the marriage certificates in the office of the Recorder of a county in California having a population of something over 100,000, half rural and half urban. In a 26-month period there were 2274 brides of all ages, including 907 or 40 per cent who were under 21 at marriage. The first births to these youthful mothers were 294 or 32 per cent of the 907. The intervals from marriage to first birth of record were tabulated in detail, enabling each reader to judge what percentage were conceived before marriage. For example, the interval was less than 8 lunar months (224 days) in 43 per cent, or less than 9 lunar months (252 days) in 56 per cent. Comparable data have been searched for in the literature and give substantial support to these high values. The causes and implications will be explored in a further paper.

REFERENCES

1. *American Child,* 1959, **41**, passim.
2. *American Med. Assn. News,* 1959, **2**, 5.
3. BAMM, P. Early Sites of Christianity. (Translated from the German original by Stanley Godman.) New York: Pantheon Books, 1957.
4. BURCHINAL, L. G., & CHANCELLOR, L. What about School-age Marriages? *Iowa Farm Sci.,* 1958, **12**, 12-14.
5. BURCHINAL, L. G. Does early dating lead to school-age marriage? *Iowa Farm Sci.,* 1959, **13**, 11-12.
6. ———. How successful are school-age marriages? *Iowa Farm Sci.,* 1959, **13**, 7-10.
7. ———. Comparisons of factors related to adjustment in pregnancy-provoked and nonpregnancy-provoked youthful marriages. *Midwest Sociol.,* 1959, **21**, 93-96.
8. ———. Adolescent role deprivation and high school marriage. *Marriage & Fam. Liv.,* 1959, **21**, 378-384.
9. ———. Are restrictive attendance policies deferents to school-age marriages. (Unpublished communication.)
10. ———. Differences in school attendance policies for married male and female students. (Mimeo unpublished.)

11. ———. Youthful marriages, problems, and programs. (Presented to the Iowa State Association of Women Deans and Counselors, Iowa State Education Association Convention, on Nov. 6, 1959, mimeo unpublished.)

12. ——— Sources of sex knowledge among Iowa high school girls. *Iowa Agricul. & Home Econom. Stat. J.* Paper No. J-3609, Iowa State University, Project 1370, no date.

13. Burgess, E. W., & Wallin, P. Engagement and Marriage. Philadelphia: Lippincott, 1953. Pp. 819.

14. Cannon, K. L. Report on study of high school girls who marry under 19 years of age. (Unpublished communication.)

15. Chambliss, R. Contributions on the vital statistics of Finland to the study of factors that induce marriage. *Amer. Sociol. Rev.*, 1957, **22**, 38-48.

16. Christensen, H. T. A comparative study of the time-interval between the marriages of parents and the birth of their first child, based on 1670 couples in Utah, 1905-1935. (Unpublished thesis, 1937.)

17. ———. Premarital pregnancy as measured by the spacing of the first birth from marriage. *Am. Sociol. Rev.*, 1953, **18**, 53-59.

18. Christensen, H. T., & Bowden, O. P. Studies in child spacing: II. The time-interval between marriage of parents and birth of their first child. *Soc. For.*, 1953, **31**, 346-351.

19. Christensen, H. T., Andrews, R., & Freiser, S. Falsification of age at marriage. *Marriage & Fam. Liv.*, 1953, **15**, 301-304.

20. Croog, S. H. Premarital pregnancies in Scandinavia and Finland. *Am. J. Sociol.*, 1952, **57**, 358-365.

21. Dublin, L. I. The Facts of Life from Birth to Death. New York: Macmillan, 1941. Pp. 461.

22. Glick, P. American Families. New York: Wiley, 1957. Pp. 240.

23. Haman, J. O. Insemination. *California Med.*, 1959, **90**, 130.

24. Henry, L. Etude statistique de l'espacement des naissances, *Population*, 1951, **6**, 425-444.

25. Hollingshead, A. B., & Redlich, F. C. Social Class and Mental Illness. New York: Wiley, 1958. Pp. 442.

26. Holm, A. *Monthly Statistical Review*, published by the City of Copenhagen, 1959, **35**, 56.

27. Kanin, E. J., & Howard, D. H. Postmarital consequences of premarital sex adjustments. *Am. Sociol. Rev.*, 1958, **23**, 557.

28. Kinsey Institute for Sex Research at Indiana University. Gebhard, P. H., Pomeroy, W. B., Martin, C. E., and Christenson, C. V. Pregnancy, Birth, and Abortion. New York: Harper, 1958. Pp. 282.

29. Landau, R. Sex, Life, and Faith. London: Faber & Faber, 1945. Pp. 319.

30. Landis, J. T., & Landis, M. G. The Marriage Handbook. New York: Prentice-Hall, 1935. Pp. 521.

31. ———. Building a Successful Marriage. Englewood Cliffs, N. J.: Prentice-Hall, 1958. Pp. 662.

32. Meaker, S. R. Sterility. *In* Davis' *Gynecology and Obstetrics*. Hagerstown, Md.: W. F. Prior Co., 1935.

33. Monahan, T. P. Does Age at Marriage Matter in Divorce? *Soc. For.*, 1953, **32**, 81-87.

34. Moss, J. J., & Gingles, R. A preliminary report on a longitudinal study of early marriage in Nebraska. (Read at the National Council of Family Relations at Eugene, Oregon, in August, 1958; unpublished.)

35. Poffenberger, S., Poffenberger, T., & Landis, J. T. Intent toward conception and the pregnancy experience. *Am. Sociol. Rev.*, 1952, **17**, 616-620.

36. Santa Barbara City Health Officer. To Dr. Helen Hart and her assistant, Mrs. Ann Kyger. (Unpublished data.)

37. Santa Barbara County Health Officer. To Dr. Joseph T. Nardo and his assistant, Mrs. Margaret Beckman. (Unpublished data.)

38. Santa Barbara County Recorder. To Mr. James G. Fowler and his assistants, Mrs. Camille Ahlman, Mrs. Lula Berggren, and Mrs. Martha Hines. (Unpublished data.)

39. SCHACHTER, J., & GRABILL, W. H. Child spacing as measured from the ages of children in the household, *Milbank Memorial Fund Quart.*, 1958, **36**, 75-85.

40. SEELEY, J. R., SIM, R. A., & LOOSLEY, E. W. Crestwood Heights. New York: Basic Books, 1956. Pp. 505.

41. Shakespeare. Romeo and Juliet, Act I, Scene ii, line 9, and Scene iii, line 17. Capulet speaks of his daughter, Juliet.

42. TIMANUS, G. L., M. D. Abortion in the United States. (Ed. by M. S. Calderone.) New York: Harper, 1958. Pp. 224.

43. TV Documentary: "I Take Thee," narrator Frank McGee, on channel 4 KRCA (NBC) in Hollywood, referred to in *TV Guide,* June 5, 1959.

44. U. S. Bureau of the Census: Vital Statistics of the U. S., 1955, *1,* LXXVI.

45. U. S. Bureau of the Census: Vital Statistics of the U. S., 1956, *1,* 6 and p. LXXXIII.

46. U. S. Bureau of the Census: Vital Statistics of the U. S., 1957, *1,* LXXXIV.

47. U. S. Bureau of the Census: *Current Population Reports, Population Characteristics,* 1958. Series P-20, No. 81, p. 3.

48. U. S. Bureau of the Census: Communication from Mr. Howard G. Brunsman, Chief of Population Division, Oct. 20, 1958.

49. WALLACE, K. M., & ODELL, E. Love Is More than Luck. An Experiment in Scientific Matchmaking. New York: Funk 1957. Pp. 237.

50. WARNER, W. L. American Life: Dream and Reality. Chicago: Univ. Chicago Press, 1953. Pp. 268.

51. WHITE, W. T., Superintendent of schools in Dallas. Communication dated May 8, 1959.

122 Olive Mill Lane
Santa Barbara, California

Published as a separate and in *The Journal of Psychology*, 1960, **50**, 399-402.

NONLINEARLY RELATED MEASUREMENT SCALES*

Kansas State Teachers College of Emporia

EUGENE S. EDGINGTON

A. INTRODUCTION

Two measurement scales are linearly related if each numerical value (measurement number) on one scale can be transformed into the corresponding value on the other scale by means of the equation $y = a + bx$, where x and y are the corresponding numerical values on the two scales and a and b are numerical constants. Conclusions based on the mean, standard deviation, product-moment correlation coefficient, or any other statistic that requires for its computation the addition or subtraction of measurement numbers, can be contradictory for any two measurement scales that are nonlinearly related (Stevens, 1951); the term *linear statistic* will be used to refer to such a statistic.

Contradictory conclusions that can arise when a linear statistic is computed with corresponding numbers from two nonlinearly related measurement scales are these: the level of significance associated with a difference between groups may be different for the two measurement scales, and the direction of a difference may be reversed; also, correlation coefficients may be different in magnitude or direction (positive or negative) for the two scales.

The term *alternative measurement scales* will be used to refer to two scales that are so related that for each measurement number on one scale there is one and only one corresponding measurement number on the other scale. Alternative measurement scales are equal in precision of measurement and contain the same amount of measurement "information."

The examples that follow will show that there are many situations in which a choice must be made between nonlinearly related alternative measurement scales. In such situations the arbitrary choice of measurement scale can affect conclusions based on linear statistics.

B. MEASUREMENT SCALES DERIVED FROM INTERDEPENDENT DIMENSIONS

In experimental situations there may be available for measurement two dimensions which are so related that one dimension absolutely determines

---

* Received in the Editorial Office on September 6, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the other. For example, if we assume circularity of the pupil of the eye, then the pupil area and pupil diameter are such dimensions. They are related by the formula $A = \frac{1}{4}\pi d^2$, so any factor that affects one dimension must affect the other. Consider a hypothetical case in which the experimenter measures both the pupil diameter and the pupil area to determine the effectiveness of pupillary conditioning. Table 1 shows the pupil sizes for two subjects, expressed in terms of pupil diameter and pupil area, under two experimental conditions.

TABLE 1
PUPIL SIZES UNDER TWO CONDITIONS

	Condition *A*		Condition *B*	
	Pupil diameter*	Pupil area**	Pupil diameter*	Pupil area**
	2.8	6.2	5.1	20.4
	7.0	38.5	5.1	20.4
	9.8	44.7	10.2	40.8
Mean:	4.9	22.3	5.1	20.4

* Millimeters.
** Square millimeters.

The pupil size in terms of pupil diameter is smaller on the average for Condition *A* than for Condition *B*, but is *larger* on the average in terms of pupil area.

One would expect conditioning of the pupil to be manifested equally well in the pupil diameter and the pupil area, and might consider the two dimensions equivalent for the purpose of testing this hypothesis, but any conclusion based on a linear statistic can be affected by the choice of dimension to measure.

C. MEASUREMENT SCALES DERIVED FROM RATIOS

Many measurement scales are derived from the ratio of measurement numbers. Convention usually determines which of the two measurement numbers will be the numerator and which the denominator, but an inversion of the fraction provides an equally good indication of the relationship between the two measurement numbers; thus, a measurement scale derived from the ratio of two measurement numbers implies the existence of a reciprocally related alternative measurement scale. Between two reciprocally related measurement scales there is a nonlinear relationship, so the computation of linear statistics for the two scales can lead to contradictory conclusions. Examples of reciprocally related scales of measurement are the following: *speed*: distance/time and time/distance; *rate of performance*: amount performed/time and time/amount performed; *density*: weight/volume and volume/weight; *pressure*: weight/surface area and surface area/weight.

In the field of chemistry two conventional ways of indicating the concentration of a solution are these: solute/solvent and solute/(solute + solvent). In the first instance there is a ratio of the solute to the rest of the material in the solution, while in the other instance there is a ratio of the solute to the total solution. These two measurement scales are nonlinearly related.

D. CONCLUSIONS

There are many situations that involve nonlinearly related alternative measurement scales; therefore, there are many situations in which conclusions based on linear statistics can be inconsistent for alternative measurement scales.

In all of the examples presented the two alternative measurement scales are ordinally related. Any statistic computed from rank order numbers instead of original measurement numbers would lead to identical conclusions for the two measurement scales; thus, rank order statistical analysis would lead to identical conclusions in many instances where the original measurement numbers would provide contradictory conclusions for alternative measurement scales.

A frequent objection to the practice of substituting rank order numbers for the original measurement numbers prior to statistical analysis is that randomness is introduced through loss of measurement information, but the preceding paragraph shows that converting to rank order numbers also can remove from the conclusions the effect of randomness resulting from an arbitrary choice between nonlinearly related alternative measurement scales.

E. SUMMARY

There are many situations in which two or more nonlinearly related alternative measurement scales are equally appropriate from the experimenter's point of view. In such situations computation of the mean, standard deviation, product-moment correlation coefficient, or any other statistic that requires for its computation the addition or subtraction of measurement numbers can lead to contradictory conclusions for the alternative measurement scales. When the alternative measurement scales are ordinally related, even though nonlinearly related, rank order statistics lead to consistent conclusions for the various measurement scales.

REFERENCE

1. STEVENS, S. S. Mathematics, measurement, and psychophysics. *In* S. S. Stevens (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 1-49.)

Department of Psychology
Kansas State Teachers College
Emporia, Kansas